



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

공학석사 학위논문

SNS 데이터를 이용한
감성분석 기반의
장소 선호도 분석기법 연구
: 서울시 테마공원을 대상으로

A Method for Analysis of Preferences of
Places based on Sentimental Analysis
using SNS Data
: Case Study on Theme Parks in Seoul,
South Korea

2017년 8월

서울대학교 대학원
건설환경공학부
채 인 영

SNS 데이터를 이용한
감성분석 기반의
장소 선호도 분석기법 연구
: 서울시 테마공원을 대상으로

A Method for Analysis of Preferences of
Places based on Sentimental Analysis
using SNS Data
: Case Study on Theme Parks in Seoul,
South Korea

지도교수 유 기 윤

이 논문을 공학석사 학위논문으로 제출함
2017년 5월

서울대학교 대학원
건설환경공학부
채 인 영

채인영의 공학석사 학위논문을 인준함
2017년 5월

위 원 장 _____ (인)

부위원장 _____ (인)

위 원 _____ (인)

국문초록

모바일 기기의 확산과 다양한 소셜 네트워크 서비스의 활성화로 사용자들은 SNS(Social Network Services)를 통해 자신의 일상과 생각을 공유하며 사회적 네트워크를 형성한다. 사용자들은 자신이 다녀온 장소에 대한 경험과 느낌을 SNS에 작성하고 이는 다른 사용자들에게 공유되며 그들의 의사결정에 영향을 미치게 된다. SNS를 통해 수많은 게시글이 업로드되면서 이로부터 의미 있는 정보를 도출하기 위한 분석이 가능하다. 따라서 본 연구에서는 SNS 텍스트를 대상으로 감성분석을 수행하고 장소 선호도를 계량화하는 산출식을 제안한다. SNS 텍스트를 이용하여 감성분석을 수행하기 위하여 기 구축된 공간 감성어 사전을 활용하였으며, SNS 텍스트의 특성을 반영하여 이모티콘을 중요한 감성정보로 고려하였다. 본 연구에서 제안한 장소 선호도 분석기법을 통해 서울에 위치한 테마공원 36곳의 선호도를 분석하였으며, 분석 결과는 향후 장소 정보 제공 서비스에서 유용하게 활용될 수 있다. 본 연구에서 제안한 기법을 적용하면 SNS 텍스트뿐 아니라 사용자들이 일상적으로 작성하는 블로그, 카페, 뉴스 기사 댓글 등의 비정형의 텍스트 데이터를 이용하여 장소 선호도를 분석할 수 있다는 점에서 의의가 있다.

주요어 : SNS 텍스트, 트위터, 이모티콘, 공간 감성어 사전,
감성분석, 장소 선호도

학 번 : 2015-22937

< 목 차 >

초 록	iii
목 차	iv
표 목차	vi
그림 목차	viii
1. 서론	1
1.1 연구 배경 및 목적	1
1.2 관련 연구	4
1.3 연구 범위 및 방법	9
2. 감성분석 기반의 장소 선호도 분석기법	13
2.1 SNS 데이터 수집 및 전처리	13
2.2 장소 POI 언급빈도수 산출	15
2.3 공간 감성어 사전 및 추가 감성어 DB를 이용한 감성분석	17
2.3.1 공간 감성어 사전을 이용한 감성분석	17
2.3.2 추가 감성어 DB를 이용한 감성분석	23
2.4 장소 선호도 산출	36
2.4.1 장소 선호도 산출식	36
2.4.2 장소 유형별 선호도 정규화	38
3. 실험 적용 및 결과	39
3.1 실험 대상 및 데이터	39
3.2 서울 테마공원 선호도 분석	41
3.2.1 서울 테마공원 POI에 대한 언급빈도수 산출	41

3.2.2 서울 테마공원 POI에 대한 감성분석	43
3.2.3 서울 테마공원 POI의 선호도 분석 결과	46
3.3 서울 테마공원 POI에 대한 감성분석 정확도 평가	50
 4. 장소 선호도 분석 활용 방안	 53
 5. 결론	 59
 참고문헌	 61
부록	66
Abstract	91

<표 목차>

표 1-1. 감성분석 관련 연구	5
표 2-1. 트위터 텍스트 메시지 예시	14
표 2-2. 불용어 제거 과정(R)	14
표 2-3. 공간 감성어 사전(일부)	18
표 2-4. 공간 감성어 사전 테이블 명세서(SSLex_single)	18
표 2-5. 공간 감성어 사전 테이블 명세서(SSLex_combi)	19
표 2-6. 감성지수벡터 및 감성값 정의	21
표 2-7. ‘올림픽공원’에 대한 감성분석 예시(공간 감성어 사전 이용) ...	22
표 2-8. 이모티콘 리스트 생성 결과	24
표 2-9. 구글맵 장소 리뷰 텍스트 일부	25
표 2-10. 추가 감성어 DB(일부)	29
표 2-11. 추가 감성어 DB 테이블 명세서	30
표 2-12. 이모티콘과 추가 감성어 DB를 이용한 감성분석 예시	30
표 2-13. 감성어와 이모티콘이 함께 사용되는 4가지 경우	31
표 2-14. 감성지수벡터 및 감성값 정의	33
표 2-15. 트위터 텍스트 감성분석 예시	34
표 2-16. ‘올림픽공원’에 대한 감성분석 예시(이모티콘 고려)	35
표 2-17. ‘올림픽공원’에 대한 예시에 따른 1차 선호도 산출	37
표 3-1. 서울 테마공원 POI	40
표 3-2. 트위터 텍스트에서 언급된 서울 테마공원 POI 및 해당 언급빈 도수	42
표 3-3. 서울 테마공원 POI를 포함하는 트위터 텍스트의 감성분석 결과 (일부)	44
표 3-4. 서울 테마공원 POI의 선호도 분석 결과	48
표 3-5. 감성분석 정확도 평가	52
표 3-6. 감성별 정확도 평가	52
표 A-1. 공간 감성어 사전 목록(SSLex_single)	67
표 A-2. 공간 감성어 사전 목록(SSLex_combi)	68
표 B-1. 추가 감성어 DB 목록	70

표 C-1. 서울 테마공원 POI를 포함하는 트위터 텍스트의 감성분석 결과 ..	77
표 D-1. 감성분석 결과에 대한 설문조사 응답 결과	85

<그림 목차>

그림 1-1. 2011년부터 2015년의 미디어 서비스 이용률 추이	1
그림 1-2. 트위터 데이터를 이용한 ‘대학’ 키워드 연관어 감성분석 결과	7
그림 1-3. 연구의 흐름도	12
그림 2-1. 장소 POI를 포함하는 SNS 텍스트 추출 과정	15
그림 2-2. 형태소 분석 예시	16
그림 2-3. 장소 POI의 언급빈도수 산출 예시	16
그림 2-4. 공간 감성어 사전을 이용한 감성분석 예시	20
그림 2-5. ‘올림픽공원’에 대한 두 감성지수벡터의 합	23
그림 2-6. 1차 추가 감성어 후보군 생성 과정	26
그림 2-7. 1차 추가 긍정 및 부정 감성어 후보군의 중복 단어 분류 과정 ..	28
그림 2-8. 트위터 텍스트에서 발견된 감성어와 이모티콘 사용의 분포	31
그림 2-9. 감성어와 이모티콘의 조합에 따른 의미 분류 결과	32
그림 3-1. 서울 테마공원 POI의 선호도 분포	49
그림 4-1. 네이버 지도의 장소 검색 결과(PC 버전)	53
그림 4-2. 네이버 지도의 ‘월드컵공원’ 검색 결과	54
그림 4-3. 네이버 지도 ‘서울 맛집’ 검색 결과(모바일 버전)	55
그림 4-4. 다음 지도의 검색결과 화면	56
그림 4-5. 다음 지도 ‘낙산공원’ 검색 결과의 평점 및 장소평	57
그림 4-6. 다음 지도 ‘낙산공원’ 검색 결과의 관련 리뷰 목록	58
그림 D-1. 감성분석 설문지 일부	84

1. 서론

1.1 연구 배경 및 목적

정보통신정책연구원의 자료에 따르면 2015년 SNS 이용률은 43.1%로, 2011년 16.8% 대비 2배 이상의 성장세를 보인다(그림 1-1). 특히, 20대의 10명 중 7명 이상 SNS를 이용하고 있을 정도로 20대의 SNS 이용률은 매우 높게 나타나며, 중장년층의 SNS 이용률 또한 증가세를 보이고 있다. 서비스별 이용률은 카카오톡, 페이스북, 트위터, 네이버밴드 순으로 나타났다.

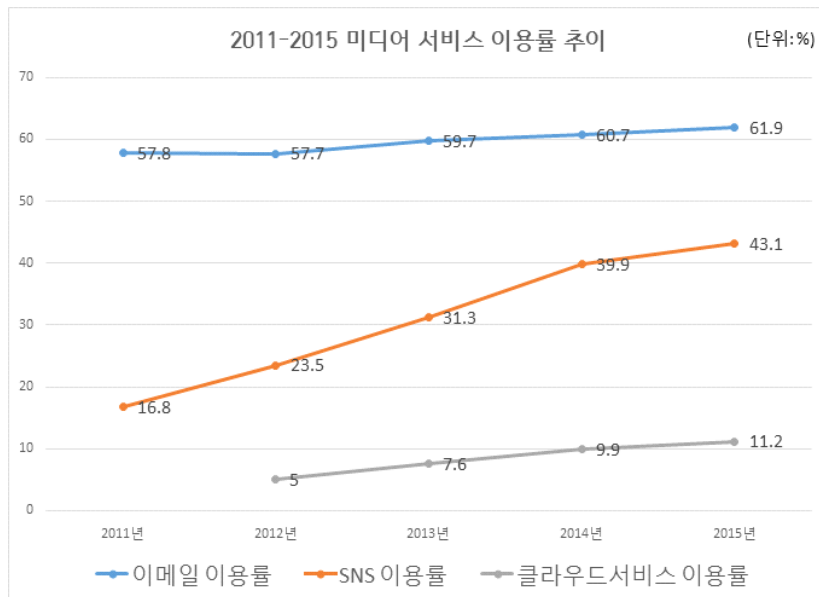


그림 1-1. 2011년부터 2015년의 미디어 서비스 이용률 추이(김윤화, 2016)

이렇듯 SNS와 인터넷의 발달 및 모바일 기기의 확산으로 인해 정보의 획득이 용이해짐에 따라 과거에 비해 특정 장소에 대한 정보에 접근할 수 있는 경로가 다양해졌으며, 장소에 대한 의견 및 정보를 생산하고

공유하는 것이 활발하게 일어나고 있다. 따라서 사람들이 SNS를 통해 다녀온 장소에 대해 소개하고 의견을 표출하며 공유하는 일상의 경험들은 다른 사용자들이 의사결정을 할 때에 중요한 정보가 될 수 있다.

그 예로 SNS 가운데 인스타그램(Instagram)의 경우, 사용자는 사진 또는 동영상과 함께 해시태그(#)를 작성하여 자신의 일상을 공유한다. 예를 들어, 사용자들이 자신이 먹은 음식 사진을 인스타그램에 업로드하면서 '#먹스타그램'과 같은 해시태그를 함께 작성하는 사례를 많이 발견할 수 있다. 실제로 '먹스타그램'은 국립국어원의 2014년 신조어로 선정되었다. 마찬가지로 사용자들은 자신이 다녀온 장소의 사진과 함께 '#여행스타그램'과 같은 해시태그를 작성하기도 한다. 이렇듯 SNS 사용이 활발해짐에 따라 이를 통해 사용자들 간의 정보 공유도 가능하다. SNS 사용자가 급증하고 수많은 게시글이 업로드되면서 방대한 양의 데이터를 분석하여 의미 있는 정보를 도출하는 것이 가능해지고, 이는 사용자들의 의사결정에 도움을 줄 수 있다.

사용자들은 모바일 기기를 이용해 일상에서의 다양한 경험을 SNS를 통해 작성한다. 트위터, 인스타그램, 페이스북과 같은 SNS는 블로그나 카페 글과 같이 PC를 갖춘 제한된 장소에서 비교적 장문의 글로 경험에 대해 자세하게 설명하는 것과 달리, 자유로운 장소와 시간에 비교적 짧고 축약된 표현으로 사용자의 경험을 작성할 수 있으므로 이는 사용자들의 경험을 담은 정보로써 분석의 가치가 있다. 따라서 본 연구는 이러한 비정형¹⁾의 SNS 텍스트를 이용하여 사용자들의 장소에 대한 선호도를 분석하고자 한다. 즉, SNS 텍스트에 나타난 장소에 대한 사용자들의 긍정 및 부정의 감성을 고려하여 사용자들로부터 언급이 많이 되고 긍정적인 감성이 표현된 장소를 분석하고자 한다. 이를 위해서는 비정형의 SNS 텍스트를 분석하여 감성정보를 추출해야한다. 따라서 본 연구는

1) 비정형 데이터(informal data)란, 형식이 정해지지 않은 데이터를 의미한다. 데이터는 형식이 정해진 정형 데이터(formal data)와 형식이 정해지지 않은 비정형 데이터로 분류된다. 페이스북, 트위터 등 SNS의 확산으로 데이터베이스에 잘 정리된 데이터가 아닌, 웹 문서, 이메일, 소셜 데이터 등 비정형 데이터가 주를 이루고 있다(한국정보통신기술협회 IT용어사전).

SNS 텍스트를 이용하여 감성분석을 수행하고 그 결과를 고려하여 장소 선호도를 산출하고자 한다.

감성분석(sentimental analysis)이란 사용자의 감성과 관련된 텍스트 정보를 자동으로 추출하는 텍스트 마이닝(text mining) 기술의 한 영역으로, 문서를 작성한 사람이 어떠한 감성을 가지고 있는가를 판단하여 분석하는 기술을 말한다(한국정보통신기술협회 IT용어사전). 주로 온라인 쇼핑몰에서 사용자의 상품평에 대한 분석이 대표적 사례로 상품에 대한 사용자의 좋고 나쁨의 감성을 판단하여 분석한다. 사용자가 작성한 리뷰는 다른 사용자의 구매 결정에 직접적인 영향을 미치기 때문에 감성분석을 통해 사용자의 반응과 요구를 파악하기 위해 주로 상품평 분석에서 감성분석이 많이 이루어지고 있다.

본 연구는 SNS 텍스트를 이용하여 감성분석을 수행하고 장소 선호도 산출식을 제안함으로써, 비정형의 텍스트 데이터로부터 장소에 대한 선호도를 도출하는 방법을 제안하고자 한다. 이는 구글맵, 다음 지도와 같이 사용자가 직접 어플리케이션이나 웹사이트에 접속하여 별점을 남기는 장소 평가와 달리 일상에서 발생하는 SNS 텍스트, 블로그 포스트, 뉴스 기사 댓글과 같은 수많은 비정형의 텍스트 데이터로부터 장소 선호도를 도출할 수 있다는 점에서 의의가 있다. 즉, 기존의 장소에 대한 별점 정보 외에 일상적으로 발생하는 수많은 SNS 텍스트, 블로그 포스트, 뉴스 기사 댓글 등을 이용함으로써 제한된 정보를 보완할 수 있다.

또한 기존 별점 정보의 경우, 5점 만점의 4.2점의 별점을 갖는 장소가 있다고 가정하였을 때, 0.8점의 부정적 평가는 무엇인지 알기 어렵다는 한계가 있다. 따라서 본 연구는 SNS 텍스트로부터 장소 선호도를 분석하고, 긍정 및 부정적 감성을 살펴봄으로써 제한적인 사용자로부터 남겨진 별점 정보 외의 장소에 대한 SNS 사용자들의 선호도를 분석할 수 있다.

최종적으로, SNS 텍스트를 이용하여 많은 사람들이 실제로 선호하는 장소를 분석함으로써 어떤 장소에 방문하고자 하는 사용자에게 장소에 대한 사전 정보를 제공할 수 있다. 이는 사용자의 의사결정에 도움을 줄 수 있으며 장소에 대한 정보를 탐색하는 데에 소모되는 시간과 비용을 절감할 수 있다.

1.2 관련 연구

감성분석 관련 연구는 주로 상품평을 대상으로 이루어졌는데, 이는 잠재적 소비자뿐만 아니라 기업의 의사결정에 중요한 자료가 되기 때문이다. 그러나 한글로 이루어진 텍스트를 분석하는 데에는 어려움이 따른다. 한글은 교착어²⁾로 영어와 달리 꾸미는 말과 어미가 복잡한 특성을 가지고 있다(장경애, 2015). 따라서 영어에 비해 감성분석 연구에 어려움이 많다. 영어권에서는 WordNet, SentiWordNet과 같은 영어의 의미어휘 목록사전을 활용하여 감성분석을 수행한 연구를 찾아볼 수 있다(Pavel, 2006; Denecke, 2008). SentiWordNet은 WordNet의 단어들 중 감성을 표현하는 어휘들을 모아 만들어진 감성어 사전으로, 각 어휘의 감성의 극성(긍정, 부정)과 극성 정도 값이 저장되어 있다(사공원, 2016).

Denecke(2008)는 Amazon의 영화 리뷰를 대상으로 SentiWordNet을 이용하여 각 리뷰로부터 감성어휘를 추출하고 해당 리뷰의 감성을 수치화하였다. 또한, Kouloumpis(2011)는 트위터에서 많이 사용되는 해시태그, ‘OMG’과 같은 약어(abbreviation), 이모티콘에 대해 데이터셋을 만들고 이를 바탕으로 감성분석을 수행하였다. Kolchyna(2015)는 트위터 데이터를 대상으로 감성어 사전 기반의 감성분석 결과를 학습 데이터로 사용하여 기계학습 알고리즘인 Naive-Bayes, SVM(Support Vector Machine)을 이용한 감성분류를 수행하였다.

한국어로 이루어진 상품평에 대한 감성분석 연구에서는 주로 의미사전을 구축하여 상품평 텍스트로부터 긍정 및 부정적 표현을 추출하고 감성을 분석하는 연구가 진행되어 왔다. 또한, 기계학습 알고리즘인 Naive-Bayes, SVM을 이용하여 상품평을 분석하거나 악성댓글을 탐지하는 연구도 이루어졌다(장경애, 2015; 홍진주, 2016).

2) 교착어: 언어의 형태적 유형의 하나. 실질적인 의미를 가진 단어 또는 어간에 문법적인 기능을 가진 요소가 차례로 결합함으로써 문장 속에서의 문법적인 역할이나 관계의 차이를 나타내는 언어로, 한국어·터키어·일본어·핀란드어가 여기에 속한다(한국정보통신기술협회 IT용어사전).

표 1-1. 감성분석 관련 연구

저자	연구주제	연구 내용	
		사전 기반	기계학습 기반
Pavel (2006)	Using WordNet for opinion mining	●	
Denecke (2008)	Using SentiWordNetfor multilingual sentiment analysis	●	
Tumasjanet et al. (2010)	Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment	●	
Kouloumpis (2011)	Twitter Sentiment Analysis: The Good the Bad and the OMG!	●	
Hridoyet et al. (2015)	Localized twitter opinion mining using sentiment analysis	●	
Kolchyna et al. (2015)	Twitter Sentiment Analysis: Lexicon Method, Machine Learning Method and Their Combination	●	●
장재영 (2009)	온라인 쇼핑몰의 상품평 자동분류를 위한 감성분석 알고리즘	●	
이윤주 외 (2014)	SNS 텍스트 콘텐츠를 활용한 오피니언마이닝 기반의 패션 트렌드 마케팅 예측 분석	●	
양민혁 외 (2014)	SNS 데이터를 활용한 국내대학 인식 및 선호도 분석	●	
임좌상 외 (2014)	한국어 트위터의 감정 분류를 위한 기계학습의 실증적 비교		●
장정애 외 (2015)	인터넷 감성기호를 이용한 긍정/부정 말뭉치구축 및 감성분류 자동화	●	●
홍진주 외 (2016)	감성분석과SVM 을 이용한 인터넷 악성댓글 탐지 기법	●	●
오평화 외 (2016)	트위터의 감정 분석을 통한 실시간 장소 추천 시스템	●	

장재영(2009)은 제품 사용자의 주관적 의견을 자동으로 분류할 수 있는 감성분석 알고리즘을 제시하였다. 상품에 대한 속성과 감성단어들에 대한 데이터베이스가 이미 구축되어 있다고 가정하고 네이트 온라인 쇼핑몰에 등록된 한글 상품평에 대해서 전체 혹은 각 속성별로 긍정 또는 부정 의견인지 판단하고 그 정도까지 계산하였다. 예를 들어, 카메라에 대한 상품평을 분석하여 카메라의 모양, 액정, 화질, 배송과 같은 속성에 대하여 감성단어를 추출하고 해당 감성의 극성과 그 정도를 나타내는 점

수를 누적하여 감성분석을 수행하였다.

장경애(2015)는 네티즌은 인터넷 쇼핑물이나 SNS 공간에서 정제된 한글이 아니라 인터넷 언어를 사용하고 있으므로 이를 고려하지 않은 상품평 분석은 정확도가 떨어진다고 보았다. 따라서 네티즌이 사용하는 인터넷 언어의 특성을 분석하여 네티즌이 감성을 표현하기 위해 사용하는 이모티콘과 한글의 초성, 특수기호로 인터넷 감성기호 사전을 구축하고 이를 통해 긍정/부정 말뭉치와 상품평의 감성분류를 자동으로 수행할 수 있는 분류방법을 제안하였다. 이 연구는 인터넷 언어의 특성을 고려하여 감성기호 사전을 구축하고 상품명에 감성기호가 포함되었을 경우 긍정 및 부정으로 분류하였으나, 해당 감성기호가 어떤 어휘와 함께 사용되었을 때에 긍정 및 부정의 의미를 갖는지 즉, 감성기호의 문맥까지는 고려하지 못했다는 한계가 있다.

상품평이 아닌 SNS 데이터를 이용하여 감성분석을 수행한 연구를 살펴보면, 양민혁(2014)은 트위터 데이터를 이용하여 국내 대학에 대한 평판을 분석하기 위하여 대학 이름과 동의어 그리고 대학과 연관된 키워드에 대한 데이터베이스를 구축하고 빈도분석, 감성분석 및 연관성 분석을 실시하여 국내 대학에 대한 인식과 선호를 분석하였다. 그 결과, ‘서울대학교’에 대한 빈도가 가장 높게 나타났으며, 연관성 분석을 통해 ‘수능, 입시, 교수, 대학생, 수시, 논술, 정시’ 등 입시와 관련된 키워드들이 대표적인 이슈로 나타났다. 또한, ‘대학’ 키워드에 대한 감성분석 결과 긍정적 반응이 81.52%로 나타났으며, 구체적인 키워드로는 ‘좋은, 합격, 희망, 최선, 열정’ 등의 단어가 나타났다(그림 1-2). 출현빈도가 가장 높게 나왔던 ‘서울대학교’를 키워드로 감성분석 결과, 부정 40%, 긍정 34%, 중립 26%의 결과가 나타났다. 부정적 의견에 ‘표절’이라는 키워드가 큰 비중을 차지하였는데, 이는 데이터 수집기간 중에 언론에 보도된 표절 문제가 대학의 이미지에도 부정적인 영향을 미치고 있다는 것을 나타낸다. 이 연구는 트위터 데이터를 이용하여 국내 대학에 대한 감성분석을 수행하였으나, 대학과 관련된 긍정 및 부정 키워드의 빈도수 즉, 출현횟수만을 이용하였으며, 이모티콘과 같은 비언어적 감성표현을 포함하는 트위터 텍스트의 특성을 고려하지 않았다는 한계가 있다.

긍부정	키워드	빈도수
긍정	좋은	14,951
부정	고민	6,990
긍정	합격	6,309
긍정	도움	6,193
긍정	희망	4,145
긍정	다양한	4,096
긍정	최선	4,073
긍정	열정	3,369
중립	세계적	2,924
부정	망하다	2,877
긍정	가고싶다	2,809
긍정	좋다	2,727
긍정	명문	2,708
긍정	성공	2,593
긍정	최고	2,464
	합계	55,927

‘대학’ 키워드 연관어 감성분석

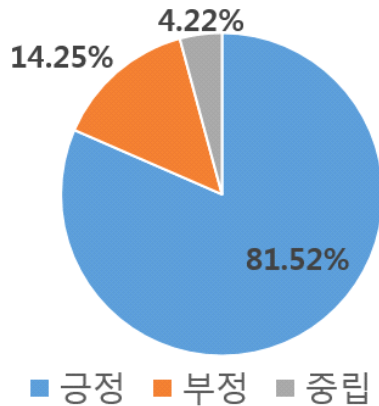


그림 1-2. 트위터 데이터를 이용한 ‘대학’ 키워드 연관어 감성분석 결과(양민혁, 2014)

오평화(2016)는 사용자의 위치 주변에서 발생한 트위터 데이터를 수집하고 감성분석을 통해 신뢰성 높은 장소를 추천하는 시스템을 제안하였다. 이를 위해 트위터에서 위치정보를 포함하는 게시글을 이용하여 형태소 분석을 한 후, 장소 추천을 위해 감성어 사전을 구축하여 긍정 및 부정을 의미하는 극성 키워드들을 정의하고 개별 트윗의 추천 점수를 도출하였다. 최종적으로 추천 점수가 높고 사용자와의 거리가 가까운 10개의 트윗을 지도 상에 추천하였다. 이 연구에서는 감성분석을 위해 박인조(2005)의 연구에서 작성된 한국어 감성단어 434개에 대하여 긍정과 부정의 극성 태그를 추가하여 감성어 사전을 구축하였다. 그러나 이는 일반적인 한국어 감성단어로써, ‘꽤쉽하다, 질투하다’와 같이 장소를 표현하는 데에 적합한 감성어가 아닌 경우가 대부분이기 때문에 장소에 대한 감성분석을 위한 적절한 감성어 사전이라고 보기 어렵다. 또한, 각 감성어에 대한 긍정 및 부정의 극성만을 태그하였을뿐, 각 감성어가 얼마나 긍정적인지 또는 얼마나 부정적인지 그 극성의 정도는 나타내고 있지 않다. 이는 결국 단순히 긍정의 단어를 포함하면 +1, 부정의 단어를 포함하면 -1의 점수를 부여하여 합산한 값을 트윗의 추천 점수로 나타내고

있어 트윗에 나타난 감성표현에 대한 극성 정도의 차이를 고려하고 있지 못하다. 또한, 트위터 텍스트의 문맥이나 트위터 텍스트에서 자주 발견되는 이모티콘 또는 한글 초성으로 이루어진 감성표현들을 감성분석의 대상으로 고려하고 있지 못하다는 한계를 지니고 있다.

임좌상(2014)은 트위터를 대상으로 형태소와 음절을 자질로 사용하여 기계학습에 따라 감성을 분류하였다. 그 결과 약 76%의 정확도로 트위터 텍스트에 포함된 감성을 분류하였다. 띄어쓰기가 불규칙적이고 ‘좋아옉ㅋㅋ’와 같은 의도적 오타가 포함되어 있는 비구조적인 트위터 텍스트의 감성을 분류하기 위해 적합한 기계학습과 분류에 영향을 미치는 자질의 정확성을 비교한 결과, SVM이 Naive-Bayes보다 감성 분류 정확성이 높게 나타났으며, 3음절 단위의 자질이 형태소 단위의 자질 보다 분류 정확성이 높게 나타났다. 이 연구는 뉴스기사와 같은 구조적인 텍스트와 달리 비구조적인 특성을 지닌 트위터 데이터에 적합한 기계학습과 자질 추출 방식을 확인하였다는 데에 의의가 있다. 그러나 트위터 텍스트에서 발견되는 이모티콘을 감성 분류를 위한 자질로 고려하지 않았다는 한계가 있다.

따라서 본 연구는 SNS 텍스트를 이용하여 감성분석을 수행하고 장소 선호도를 분석하기 위하여 장소를 표현하는 데에 적합한 감성어로 이루어진 공간 감성어 사전을 활용하고, SNS 텍스트의 특성을 고려하기 위하여 한글 초성 및 이모티콘을 감성표현으로 간주하여 감성분석의 대상으로 고려하고자 한다.

1.3 연구 범위 및 방법

사람들은 특정 장소에 방문하고 그곳에서 느낀 감성과 경험을 SNS를 통해 공유한다. 이는 참여형 지리정보(Volunteered Geographic Information, VGI)에 속한다고 볼 수 있으며, 텍스트, 사진, 동영상 등으로 표현될 수 있다. 본 연구의 대상은 비정형의 SNS 텍스트이며, 이로부터 특정 장소에 대한 선호도 정보를 도출한다면 해당 POI(Point Of Interest)의 위치정보뿐 아니라 정성적인 속성이라고 할 수 있는 사람으로부터 선호되는 정도를 정량적으로 계량화(quantification)하여 새로운 속성정보로 활용이 가능하다.

장소 선호도는 장소를 이용하는 사람이 해당 장소에 대해 가지게 되는 좋아하는 정도이다. 어떤 장소에 대한 선호도는 그 장소와 사람 간의 다양한 상호작용을 통해서 형성된다. 즉, 이용자의 다양한 활동과 상태, 장소의 상태와 상황들이 복합적으로 작용하여 발생하는 것이다. 이와 같은 복합적인 작용이 특정 장소에 대한 특별한 이미지를 갖게 하고 특별한 정체성을 형성하여 이용자에게 인식된다(어정연, 2013).

본 연구에서 장소 선호도 분석을 위해 이용하는 SNS 텍스트는 트위터 텍스트이며, 트위터는 140자 이내 단문으로 개인의 의견이나 생각을 공유하고 소통하는 소셜 네트워킹 및 마이크로 블로그 서비스이다. twitter(지저귀다)의 뜻 그대로 재잘거리듯이 일상의 작은 얘기들을 수시로 짧게 남길 수 있는 온라인 공간을 의미한다.

본 연구에서 트위터 텍스트를 이용하여 장소 선호도를 분석하는 방법은 그림 1-3과 같다.

첫 번째 단계에서는 SNS 텍스트, 본 연구에서는 트위터 텍스트를 오픈 API를 이용하여 수집하고 텍스트 분석을 위한 전처리 과정을 거친다. 이는 트위터 텍스트에 존재하는 URL, RT, @, #, 영문자, 숫자와 같은 불용어(stopwords)를 제거하는 과정을 의미한다.

두 번째 단계에서는 선호도를 분석하고자 하는 장소의 유형에 따라 POI의 리스트를 생성하고, 이를 포함하는 SNS 텍스트를 추출한다. 이를

대상으로 자연어 처리를 위해 형태소 분석을 수행하고, 추출된 전체 SNS 텍스트 내에서 각 장소 POI의 언급빈도수를 산출한다. 이는 각 장소 POI에 대한 SNS 사용자들의 방문 및 관심 정도를 반영한다.

세 번째 단계에서는 공간 감성어 사전 및 이모티콘을 고려하여 감성분석을 수행한다. 장소에 대한 감성분석을 위해서는 장소를 표현할 때 사용하는 적절한 감성어휘 정보가 필요하다. 따라서 본 연구에서는 기 구축된 공간 감성어 사전(이영민, 2017)을 이용한다. 공간 감성어 사전은 장소를 나타내기 위해 적합한 공간 감성어와 해당 감성의 극성(긍정, 중립, 부정) 그리고 확률점수(0~1)를 저장하고 있다.

또한, 트위터 텍스트의 특성을 고려하기 위하여 이모티콘을 감성기호로서 중요한 감성정보로 고려한다. 이모티콘이란 컴퓨터 자판의 문자, 기호, 숫자 등을 조합해 감성이나 의사를 전달할 때 사용하는 하나의 언어이다. 감성을 뜻하는 ‘이모션(emotion)’과 그림 기호를 뜻하는 ‘아이콘(icon)’을 합친 말로, 1980년대 초 미국의 카네기 멜론 대학에서 컴퓨터 자판을 이용해 ‘:-)’ 같은 모양을 사용하면서 보급되기 시작하였다³⁾. 한글로 이루어진 텍스트 분석을 위해서는 한글을 제외한 이모티콘, 특수기호, 영문자 등을 분석의 의미가 없다고 판단하여 제거하는 것이 보편적이지만, 본 연구에서는 이모티콘을 중요한 감성정보로 여겨 불용어 제거 단계에서 제거하지 않고 텍스트에 남겨둔다.

본 연구에서는 선행 연구(장경애, 2015)에서 정의한 인터넷 감성기호 일부와 연구자가 트위터 텍스트를 관찰한 결과 트위터 사용자가 자주 사용하는 이모티콘을 정의하여 이모티콘 리스트를 생성한다. 또한, 문맥을 고려하기 위하여 구글맵 장소 리뷰 데이터를 활용하여 어떤 단어와 어떤 이모티콘이 함께 사용되었을 때, 긍정 또는 부정의 의미를 갖는지 파악하고자 한다.

마지막 단계에서는 앞서 도출한 각 장소 POI에 대한 언급빈도수와 감성분석 결과를 종합적으로 고려하여 본 연구에서 제안하는 장소 선호도 산출식에 의하여 각 장소 POI에 대한 1차 선호도를 산출한다. 이는 특정

3) <https://ko.wikipedia.org/wiki/이모티콘> (접속일: 2017년 6월 10일)

장소 POI에 대해 SNS 사용자들이 많이 방문하고 관심이 높을수록 언급 빈도수가 높고, 장소에 대한 긍정적인 의견과 감성이 많이 나타날수록 긍정의 감성점수가 높게 나타나므로 해당 장소 POI의 선호도는 높게 나타나는 개념이다.

최종적으로 해당 장소 유형에 속하는 여러 POI 가운데 가장 선호도가 높은 POI부터 가장 선호도가 낮은 POI까지 도출된다. 장소 유형에 따라 해당 POI들의 1차 선호도 값의 범위가 다르기 때문에 이들 간의 비교를 위해서 정규화(normalization)할 필요가 있다. 따라서 본 연구에서 제안하는 장소 선호도 산출식에 의하여 산출된 1차 선호도를 정규화함으로써 최종 장소 선호도를 도출할 수 있다.

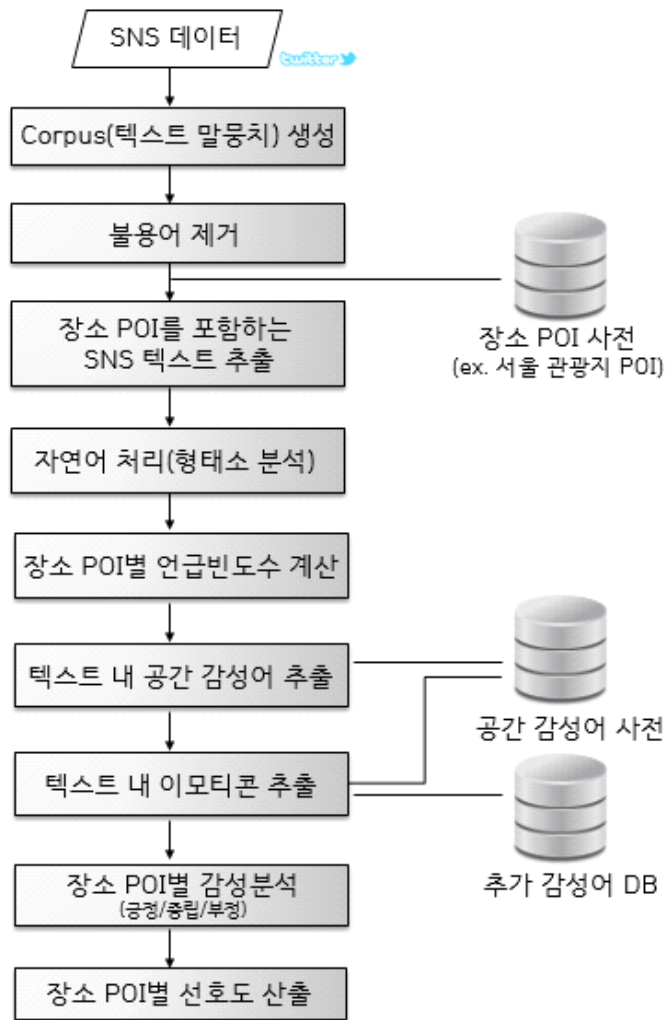


그림 1-3. 연구의 흐름도

2. 감성분석 기반의 장소 선호도 분석기법

2.1 SNS 데이터 수집 및 전처리

대부분의 SNS는 각자의 플랫폼을 이용한 서드파티(third party)⁴⁾들의 다양한 서비스 개발을 독려하기 위하여 오픈 API(Application Programming Interface)⁵⁾를 제공한다. 본 연구에서 이용하는 SNS 데이터인 트위터 데이터는 REST(Representational State Transfer) 기반의 오픈 API를 이용하여 수집할 수 있다. 본 연구에서는 일정 기간 동안 발생한 트위터 데이터를 수집하고 데이터의 다양한 속성 정보(사용자 ID, 텍스트, 작성 시간 등) 필드 값이 모두 동일한 경우, 중복된 데이터로 판단하여 제거한다. 그 후, 트위터 데이터의 다양한 속성 정보 중 트위터 텍스트 정보가 담긴 'text' 필드를 본 연구에서 사용한다. 텍스트 데이터 처리를 위하여 본 연구에서 사용하는 프로그래밍 언어는 R x64 3.2.5이며, 한글로 이루어진 자연어의 처리를 위해 KoNLP(Korean Natural Language Processing) 패키지를 이용한다.

수집된 트위터 데이터 가운데 텍스트를 리스트 형태로 저장하여 말뭉치를 생성한다. 다음으로 분석에 의미가 없다고 판단되는 불용어를 제거하기 위해서 트위터 텍스트의 구성요소를 살펴볼 필요가 있다.

트위터 텍스트는 표 2-1과 같이 오로지 한글로만 이루어져 있지 않으며, 트위터의 기능에 따라 약자(abbreviation) 또는 URL의 형태로 영문자를 포함하고 있다. 메시지 앞에 'RT'가 붙어있으면 이는 다른 사용자로부터 전달된 메시지(retweet)란 뜻이다. '#'은 해시태그(hashtag)로 사

4) 서드파티: 하드웨어나 소프트웨어 등의 제품을 제조하고 있는 메이커나 그 계열 회사 또는 기술 제휴를 하고 있는 기업 이외의 기업을 총칭한다(전자용어사전).

5) 오픈 API: 인터넷 이용자가 일방적으로 웹 검색 결과 및 사용자인터페이스(UI) 등을 제공받는 데 그치지 않고 직접 응용 프로그램과 서비스를 개발할 수 있도록 공개된 API를 말한다(매일경제용어사전).

용자가 트윗을 작성할 때, 주요 키워드 앞에 ‘#’을 추가할 수 있으며, 이는 검색어 기능을 할 수 있다. ‘@사용자 ID’는 멘션(mention)으로 제3자를 언급하는 것을 의미한다. ‘https://’로 시작하는 영문 URL은 해당 트윗의 출처를 의미한다(남민지, 2015).

표 2-1. 트위터 텍스트 메시지 예시

RT @lghopper From Stage...마지막엔? #장기하와얼굴들 #gmf2015 <https://t.co/jFR0mBLZJy>

따라서 ‘RT’, ‘#’, ‘@’과 URL 주소 그리고 숫자 및 영문자는 추후 트위터 텍스트의 감성분석에 있어 의미를 가진 요소가 아니므로 불용어 제거 과정을 거쳐야한다. 이를 위해서 R에서 제공하는 gsub 함수를 이용한다. 표 2-2는 R에서 해당 불용어를 공백으로 대체하여 불용어를 제거하는 과정이다.

표 2-2. 불용어 제거 과정(R)

```
ntext<-gsub("http[s]?://(?:[a-zA-Z][0-9]|[$-_@.&+]|[*\\(\\)])|(?:%[0-9a-fA-F][0-9a-fA-F]))+", ' ', ntext) # url 제거
ntext<-gsub('[A-Za-z0-9]', ' ', ntext) # 영어, 숫자 제거
ntext<-gsub('#', ' ', ntext) # 해시태그(#) 제거
ntext<-gsub('@', ' ', ntext) # @ 제거
```

불용어 제거 과정에서 중요한 점은 보편적인 불용어 제거 단계에서 제거하는 한글 초성 및 특수문자 기호는 제거하지 않고 남겨둔다는 점이다. 본 연구에서는 ‘ㄱㄱ’, ‘ㅍㅍ’, ‘^^’ 와 같은 한글 초성 및 특수문자 기호를 감성표현으로서 감성분석의 대상으로 고려하기 때문이다. 실제로 많은 사용자가 입력 공간이 협소한 스마트 폰으로 트윗을 작성하기 때문에 이모티콘을 포함하여 작성하는 경우가 많다(홍초희, 2012).

NC는 명사, JC는 격조사, II는 감탄사, PA는 형용사, EF는 종결어미를 뜻하는 품사 태그이다.

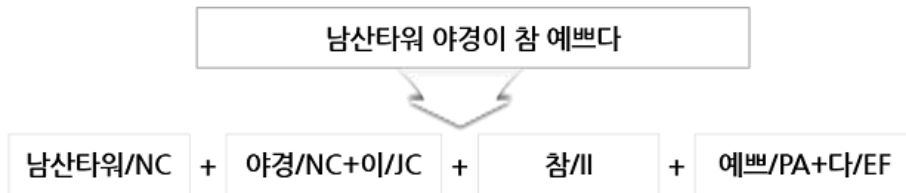


그림 2-2. 형태소 분석 예시

형태소 분석이 이루어진 트위터 텍스트를 대상으로 각 장소 POI의 언급빈도수(TF, Term Frequency), 즉, 전체 트위터 텍스트 내에 나타나는 해당 단어(각 장소 POI)의 총 빈도수를 산출한다. 이는 SNS 내에서 특정 장소 POI의 언급이 많이 될수록 그 장소 POI에 대하여 SNS 사용자들의 방문 및 관심 정도가 높다고 가정하는 것이다. 예를 들어, ‘롯데월드’가 포함된 트위터 텍스트가 115개, ‘여의도 한강공원’이 포함된 트위터 텍스트가 64개, ‘남산타워’가 포함된 트위터 텍스트가 41개라면, 이는 곧 각 장소 POI에 대한 언급빈도수가 된다(그림 2-3).



그림 2-3. 장소 POI의 언급빈도수 산출 예시

2.3 공간 감성어 사전 및 추가 감성어 DB를 이용한 감성분석

2.3.1 공간 감성어 사전을 이용한 감성분석

각 장소 POI에 대한 SNS 사용자들의 의견 및 감성을 살펴보기 위하여 트위터 텍스트에 대한 감성분석을 실시한다. 이는 장소에 대한 감성을 나타내는 표현인 공간 감성어가 저장된 사전을 활용하여 이루어질 수 있다. 따라서 본 연구에서는 이영민(2017)에 의해 구축된 공간 감성어 사전(Spatial Sentiment Lexicon)을 활용한다. 이영민(2017)에 따르면 공간 감성어 사전은 장소를 설명할 때 사용되는 감성어휘에 대한 데이터베이스를 의미하며, 여러 종류의 장소들 중에서 ‘테마공원’을 대상으로 하였다. 또한, 공간 감성어 사전 구축에 있어 다음과 같이 네 가지 측면을 고려하였다.

- 1) 감성어는 감성의 극성과 극성의 정도에 대한 수치 값을 보유해야 한다.
- 2) 일부 단어들은 함께 사용되는 서술어에 따라 감성의 극성이 달라질 수 있으므로 이를 고려해야 한다.
- 3) 공간 감성어 사전이 감성분석에 이용될 때 글의 문맥을 반영할 수 있도록 지원해야 한다.
- 4) 장소의 특성에 따라 달라지는 다양한 감성어를 고려해야 한다.

이영민(2017)에 의해 구축된 공간 감성어 사전의 일부와 해당 테이블 명세서는 표 2-3, 표 2-4, 표 2-5와 같으며, 공간 감성어 사전의 전체 목록은 <부록 A>에 수록하였다. 이는 독립적인 하나의 감성어휘 114개로 이루어진 SSLex_single 사전과 사람들이 장소에 대한 리뷰글을 작성할 때 자주 사용하는 두 어휘의 조합 104개를 저장한 SSLex_combi 사전으로 이루어져있다.

표 2-3. 공간 감성어 사전(일부)

Table Name	SSLex_single			Table Name	SSLex_combi			
ID	SP	Pt	Pb	ID	SS	SP	Pt	Pb
S_1	가깝/PA	1	1.00000	C_1	건물/NC	많/PA	0	0.76351
S_2	넓/PA	1	0.98824	C_2	나무/NC	적/PA	-1	0.98925
S_3	추천/NC	1	1.00000	C_3	사람/NC	많/PA	0	0.68456
S_4	괜찮/PA	1	0.98374	C_4	주차장/NC	적/PA	-1	1.00000
S_5	나쁘/PA	-1	0.99213	C_5	분위기/NC	색다르/PA	1	0.98113
S_6	봄비/PA	0	0.47651	C_6	야경/NC	예쁘/PA	1	1.00000
S_7	비좁/PA	-1	0.99270	C_7	물가/NC	비싸/PA	-1	0.96774
S_8	건/PV	0	0.66000	C_8	경치/NC	좋/PA	1	1.00000
S_9	최고/NC	1	1.00000	C_9	여유/NC	느끼/PV	0	0.79412
S_10	힘들/PV	-1	1.00000	C_10	여유/NC	즐거/PV	1	1.00000
...

표 2-4. 공간 감성어 사전 테이블 명세서(SSLex_single)

Table Name	SSLex_single				
Column Name	Data Type	Nullable	PK	FK	Description
ID	CHAR(10)	NOT NULL	PK		일련번호
SP	VARCHAR(50)	NOT NULL			공간 서술어(spatial predicate) 어근(root)과 품사(NC: 명사, PA: 형용사, PV: 동사)가 결합된 형태
Pt	INT	NOT NULL			감성의 극성(polarity) (긍정: 1, 부정: -1, 중립: 0)
Pb	FLOAT(7, 6)	NOT NULL			감성확률(probability) 극성의 정도에 대한 수치값(0~1)

표 2-5. 공간 감성어 사전 테이블 명세서(SSLex_combi)

Table Name	SSLex_combi				
Column Name	Data Type	Nullable	PK	FK	Description
ID	CHAR(10)	NOT NULL	PK		일련번호
SS	VARCHAR(50)	NOT NULL			공간 주제어(spatial subject) 어근(root)과 품사(NC: 명사, PA: 형용사, PV: 동사)가 결합된 형태
SP	VARCHAR(50)	NOT NULL			공간 서술어(spatial predicate) 어근(root)과 품사(NC: 명사, PA: 형용사, PV: 동사)가 결합된 형태
Pt	INT	NOT NULL			감성의 극성(polarity) (긍정: 1, 부정: -1, 중립: 0)
Pb	FLOAT(7, 6)	NOT NULL			감성확률(probability) 극성의 정도에 대한 수치값(0~1)

앞서 형태소 분석이 이루어진 장소 POI를 포함하는 트위터 텍스트와 공간 감성어 사전에 저장된 공간 감성어의 키워드 매칭을 통해 텍스트 내 감성정보를 추출한다. 그 과정에서 그림 2-4와 같이 매칭되는 감성어를 기준으로 ± 3 어절(word phrase) 거리 내에 장소 POI, 지시대명사인 ‘여기’, ‘이곳’ 또는 장소의 속성을 나타내는 공간 주제어(SS, Spatial Subject, (예) 야경, 분위기, 경치)가 위치하였을 경우에 문맥상 장소에 대한 감성표현으로 판단하여 해당 감성어를 추출한다.

이는 이종혁(2014)의 연구에서 제안하는 어휘 수식 거리를 단문인 트위터의 특성에 맞게 변경한 것이다. 이종혁(2014)은 상품에 대한 의견이 담긴 블로그 글을 대상으로 문맥을 고려한 감성분석을 수행하기 위하여 어휘 수식 거리를 적용하였는데, ‘명동칼국수는 정말 맛있다’라는 문장에서 주어인 ‘명동칼국수’와 서술어인 ‘맛있다’ 사이에 감성을 나타내는 관형어와 부사가 추가될수록 문장이 길어지고 감성의 깊이가 깊어진다고 보았다. 따라서 주어와 감성 서술어 사이의 단어 수를 감성 깊이의 가중치로 고려하였다. 단문인 트위터의 경우, 블로그와 달리 사용자들이 텍스트를 작성할 때, 주어, 목적어, 서술어의 완결된 구조의 문장을 갖추어

작성하는 경우가 많지 않으므로, 주어와 서술어 사이의 거리가 멀어질수록 오히려 지칭하는 의미가 모호해지거나 부정확할 우려가 있다. 따라서 주어, 목적어, 관형어 또는 부사, 서술어의 문장 구조를 갖추었을 경우, 주어와 서술어 간의 어휘 수식 거리인 3을 기준으로 장소에 대한 표현을 포착하고자 하였다.

다음으로, 각 트위터 텍스트별로 추출된 공간 감성어의 해당 감성의 극성(Pt)과 감성확률(Pb)을 곱한 값을 감성값이라 하고, 이 값이 0보다 크면 긍정값(PT, Positive index of Text)으로, 0보다 작으면 부정값(NT, Negative index of Text)으로 구분하여 각 트위터 텍스트별로 누적하여 합산한다(그림 2-4).

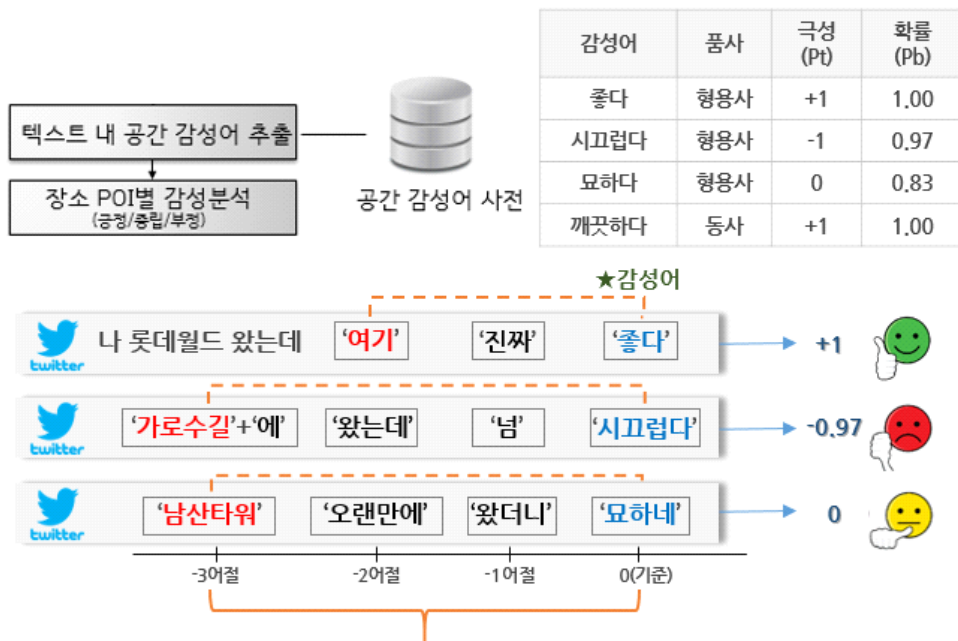


그림 2-4. 공간 감성어 사전을 이용한 감성분석 예시

모든 트위터 텍스트에 대하여 공간 감성어 사전과의 키워드 매칭 과정을 거쳐 위에서 언급한 경우에 해당하는 경우 긍정값과 부정값이 부여된다. 이 결과를 바탕으로 각 장소 POI에 대하여 해당 트위터 텍스트들의 긍정값과 부정값을 각각 합산하여 이를 해당 장소 POI에 대한 긍정지수(S_P)와 부정지수(S_N)로 정의한다. 여기서, 이 두 감성지수는 감성의 방향과 크기를 지니므로 1차원 벡터로 표현이 가능하다. 표 2-6은 이를 수식으로 정의한 것이다. 이 두 감성지수벡터의 크기비교를 통해 해당 장소 POI에 대한 긍정적 감성과 부정적 감성의 우세를 파악할 수 있다.

표 2-6. 감성지수벡터 및 감성값 정의

	장소 POI에 대한 감성지수벡터(\vec{S})	텍스트의 감성값(T)
긍정 (Positive)	$\vec{S}_P = \sum_{i=1}^{F_k} PT_{(i)}$	$PT_{(i)} = \sum_{j=1}^n \{match(i,j) \times Pb_{(j)}\}$
부정 (Negative)	$\vec{S}_N = \sum_{i=1}^{F_k} NT_{(i)}$	$NT_{(i)} = - \sum_{j=1}^n \{match(i,j) \times Pb_{(j)}\}$

여기서, i 는 트위터 텍스트, k 는 장소 POI, F_k 는 k 의 언급빈도수, $Pb_{(j)}$ 는 감성어 j 의 감성확률, n 은 i 에 존재하는 j 의 개수를 의미한다. 또한, $match(i,j)$ 는 아래와 같이 경우에 따라 0 또는 1의 값을 가진다.

$$match(i,j) = \begin{cases} 1 & : \text{텍스트 } i \text{에 포함된 } j(\text{형용사, 동사, 명사}) \text{가 공간 감성어 사전에} \\ & \text{존재하고, '장소 POI/ 지시대명사/공간주제어'와의 거리가} \\ & \text{3어절 이내일 경우} \\ 0 & : \text{그 외의 경우} \end{cases}$$

표 2-7은 ‘올림픽공원’에 대한 7개의 트위터 텍스트에 대해 감성분석한 것으로, 긍정지수벡터(\vec{S}_P)와 부정지수벡터(\vec{S}_N)를 산출하고, 이 두 벡터의 크기비교를 통해 ‘올림픽공원’에 대한 긍정적 감성이 더 많이 나타났음을 알 수 있다. 즉, 두 벡터의 합이 긍정(+)이므로 긍정적 감성이 우세함을 알 수 있다. 그림 2-5은 이를 나타낸 것이다.

표 2-7. ‘올림픽공원’에 대한 감성분석 예시(공간 감성어 사전 이용)

i (번호)	Text (i)	감성어 추출 및 개별 텍스트의 감성값 산출	
		$PT(i)$ (Positive index of Text(i))	$NT(i)$ (Negative index of Text(i))
1	급 번개 [나들이 올림픽공원] (긍정(+1), 1.0)	1.0	NA
2	올림픽공원 수영장 다녀오다가 하늘이 맑아 찍어본 하늘사진	NA	NA
3	[주차장도 넓]고 [주차비용도 싸]서 [데이트하러 오기 좋다] (긍정(+1), 0.98) (긍정(+1), 0.89) (긍정(+1), 1.0)	0.98+ 0.89+1.0 = 2.87	NA
4	오늘은 회사언니랑 [올림픽공원까지 산책]하러 (긍정(+1), 1.0)	1.0	NA
5	[사람 많]고 너무 [힘들었던 하루 올림픽공원] (중립(0), 0.685) (부정(-1), 1.0)	NA	-1.0
6	올림픽공원 행사하는지 [차가 너무 많다].그래도 [단풍은 예쁘네] (부정(-1), 0.89) (긍정(+1), 1.0)	1.0	-0.89
7	[아이들] 데리고 왔더니 좋아하네 자주 와야지 올림픽공원 (긍정(+1), 1.0)	1.0	NA
빨강: 부정 초록: 중립 파랑: 긍정 $\sum_{i=1}^{F_k} PT(i) \text{ or } NT(i)$		+6.87 \vec{S}_P	-1.89 \vec{S}_N

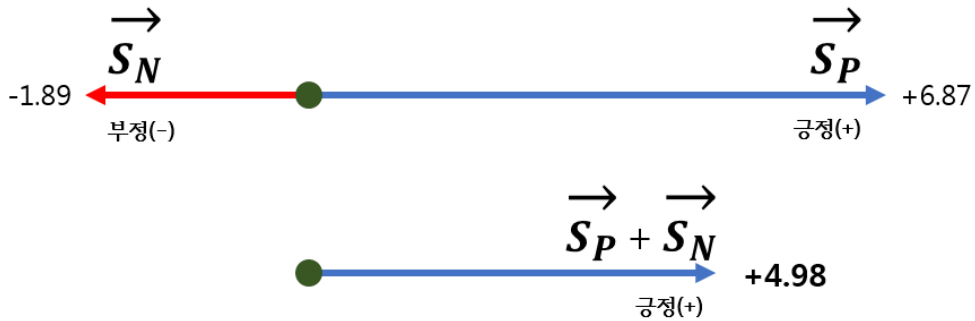


그림 2-5. ‘올림픽공원’에 대한 두 감성지수벡터의 합

2.3.2 추가 감성어 DB를 이용한 감성분석

본 연구에서는 기 구축된 공간 감성어 사전을 활용함과 더불어 이모티콘을 중요한 감성정보로 고려하여 트위터 텍스트의 특성을 반영한 감성 분석을 수행하고자 한다. 이를 위해 불용어 제거 단계에서 한글 초성 및 특수문자 기호를 제거하지 않는다. 본 연구에서는 선행 연구(장경애, 2015)에서 정의한 인터넷 감성기호 일부와 연구자가 트위터 텍스트를 관찰한 결과 트위터 텍스트에서 발견되는 주요 이모티콘을 정의하여 이모티콘 리스트를 생성하였으며, 그 결과는 표 2-8과 같다.

또한, 표 2-8에서 정의한 이모티콘을 바탕으로 문맥을 고려하기 위하여 구글맵 장소 리뷰 데이터를 활용하여 어떤 단어와 어떤 이모티콘이 함께 사용되었을 때, 긍정 또는 부정의 의미를 갖는지 파악하고자 한다. 예를 들어, ‘좋아요ㅠㅠ’의 경우 ‘ㅠㅠ’는 부정의 이모티콘이지만 ‘좋아요’의 긍정적 의미를 강조하는 역할을 한다. 반면, ‘진짜 별로네^^’의 경우 ‘^^’는 긍정의 이모티콘이지만 ‘별로네’라는 부정적 의미와 함께 비꼬는 역할을 한다.

표 2-8. 이모티콘 리스트 생성 결과

감성	이모티콘	비고
긍정(+1)	^^	선행연구 참고
	ㄱㄱ	
	ㅎㅎ	
	:)	트위터 관찰 결과
	^^	
	♥	
	♡	
	~	
	+ _ +	
부정(-1)	ㅍㅍ	선행연구 참고
	ㅌ	
	_ _;;	
	_ _	트위터 관찰 결과
	ㅡㅡ	
	ㅌ.ㅌ	
	ㅍㅍ	
	:(
	...	

따라서 본 연구에서는 공간 감성어 사전의 제한된 감성어 개수의 한계를 극복하고자 구글맵 장소 리뷰 데이터를 이용하여 이모티콘과 함께 사용되었을 때 의미를 갖는 단어의 집합인 추가 감성어 DB를 생성하였다. 구글맵 장소 리뷰 데이터를 이용한 이유는 리뷰의 별점 정보를 통해 긍정 및 부정의 리뷰로 분류가 가능하여 긍정적인 의미로 사용되는 확률이 높은 감성어와 부정적인 의미로 사용되는 확률이 높은 감성어를 도출하는 것이 가능하기 때문이다.

이를 위해 먼저, 구글 API를 이용하여 구글맵의 ‘테마공원’ POI 검색 결과에 따른 사용자들이 남긴 장소 리뷰데이터를 수집하였다. 구글 API의 정책상 하나의 POI당 최대 5개의 리뷰를 수집할 수 있으며, 그 결과 별점 1점부터 5점까지의 리뷰 총 7,091건을 수집하였다. 이를 이용하여 긍정의 감성어와 부정의 감성어를 추출하기 위하여 사용자가 별점 5점을 남긴 리뷰(긍정)와 별점 1점을 남긴 리뷰(부정)만을 이용하였다. 표 2-9는 별점 5점과 별점 1점을 받은 구글맵 장소 리뷰 데이터의 일부이다.

표 2-9. 구글맵 장소 리뷰 텍스트 일부

별점	텍스트	별점	텍스트
5점	정동길 걷기도 좋고 주위 풍경이 너무 좋아요	1점	번잡하고 주차하기 힘들고 들어가는 순간 밖에서 사진 찍고 집에 갔으면 되었는데 하는 생각이 들었다
	큰 나무가 많아 숲속 깊은 곳에 온 느낌이 들어요		기대하고 갔다가 실망 너무 관리가 안되고 동물들이 너무 아파보였음 ㅜㅜ 겨울에만 그런건지 블로그에서 보던 평가와는 조금 많이 다름
	서울의 허파와 같은 남산. 사시사철 눈이오나 비가오나 한결같이 아름답고 고마운 산.		볼게 없고 재미도 없고 별도 많고 지저분함. 참고로 다리도 아프고 그저 그렇다.
	항상 다양한 전시가 열려 언제나 편하게 방문하기 좋은 곳입니다.		좁고 불친절
	매우 깨끗하고 편리한 도시 속 캠핑장 샤워장은 오전은 9시까지만 운영		너무 멀고 잘못 들어오기 쉬워서 안 좋다

이를 이용하여 1차 추가 감성어 후보군을 생성하기 위한 과정은 그림 2-6과 같다. 먼저, 별점 5점의 구글맵 장소 리뷰 텍스트 3,040개를 이용하여 형태소 분석을 하고 명사(NC), 형용사(PA), 동사(PV) 품사 태그가 붙은 단어를 추출하였다. 이는 명사(NC), 형용사(PA), 동사(PV)가 실질적 의미를 나타내는 어근이기 때문이다(이영민, 2017). 그 결과, 5,983개의 단어가 추출되었으며 이들의 출현횟수의 평균(M)을 산출하였다. 5,983개의 단어들의 평균 출현횟수(M)는 8.5008회(약 9회)로, 전체 단어 가운데 9회 이상의 출현횟수를 가지는 단어를 추출하였다. 이는 별점 5점의 긍정 리뷰 문서 내에서 평균 이하로 사용되는 단어들을 걸러내기 위함이다. 다음으로, 추출된 단어들 가운데 단어의 길이가 6자 이상인 단어를 제외하였다. 이는 국립국어원이 발표한 자주 쓰이는 한국어 기초 낱말⁶⁾ 가운데 가장 긴 단어가 ‘즐거워하다, 찾아다니다’와 같은 5자의 단

6) https://ko.wiktionary.org/wiki/부록:자주_쓰이는_한국어_낱말_5800
(접속일: 2017년 6월 10일)

어이므로, 이 이상의 길이를 가진 단어는 띄어쓰기가 잘 되어있지 않은 리뷰 텍스트들의 부정확한 형태소 분석 결과로 보고 제거하기 위함이다. 예를 들어, ‘새벽에무료로산책다녀왔는데정말좋더군/NC’, ‘정말훌륭하답니/NC’와 같은 단어는 잘못된 형태소 분석의 결과로 도출된 6자 이상의 길이를 가진 단어들이므로 제거되었다. 결과적으로 추가 감성어의 단어는 1자~5자의 길이를 가진 단어들로 구성되며, 463개의 단어가 도출되었다. 마지막으로, 이 단어들 가운데 공간 감성어 사전에 저장된 긍정의 감성어와 중복되는 항목을 제거하여 최종적으로 432개의 1차 추가 긍정 감성어 후보군을 생성하였다.

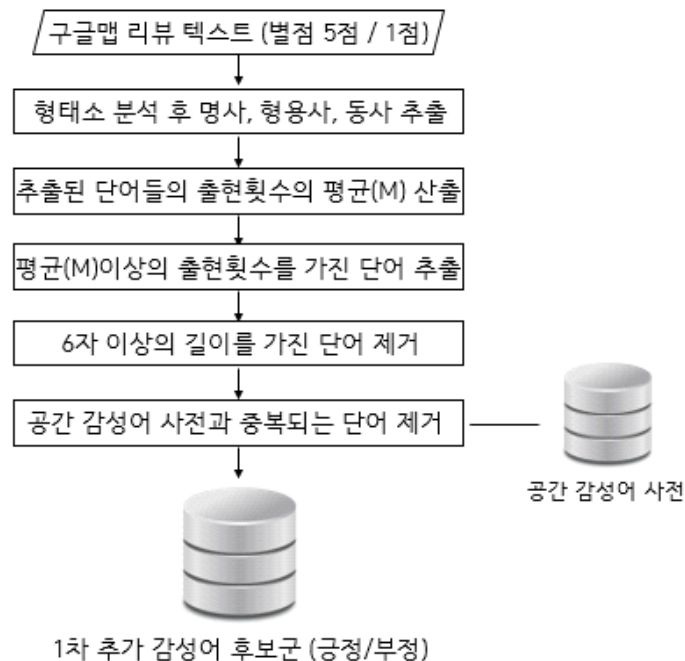


그림 2-6. 1차 추가 감성어 후보군 생성 과정

마찬가지로, 별점 1점의 구글맵 장소 리뷰 텍스트 375개를 이용하여 동일한 과정을 수행하였다. 추출된 명사(NC), 형용사(PA), 동사(PV)의 단어 1,590개의 평균 출현횟수(M)는 2.0968회(약 2회)이므로, 이들 가운데 2회 이상의 출현횟수를 가지는 단어를 추출하였다. 다음으로, 6자 이

상의 길이를 가진 단어를 제거한 결과, 397개의 단어가 추출되었으며, 이 단어들 가운데 공간 감성어 사전에 저장된 부정의 감성어와 중복되는 항목을 제거하여 최종적으로 388개의 1차 추가 부정 감성어 후보군을 생성하였다.

지금까지 생성한 1차 추가 긍정 감성어 후보군 432개의 단어와 1차 추가 부정 감성어 후보군 388개의 단어 가운데 서로 중복되는 단어가 존재하는지 비교하였다(그림 2-7). 그 결과, 97개의 단어가 1차 추가 긍정 및 부정 감성어 후보군에 모두 존재하였다.

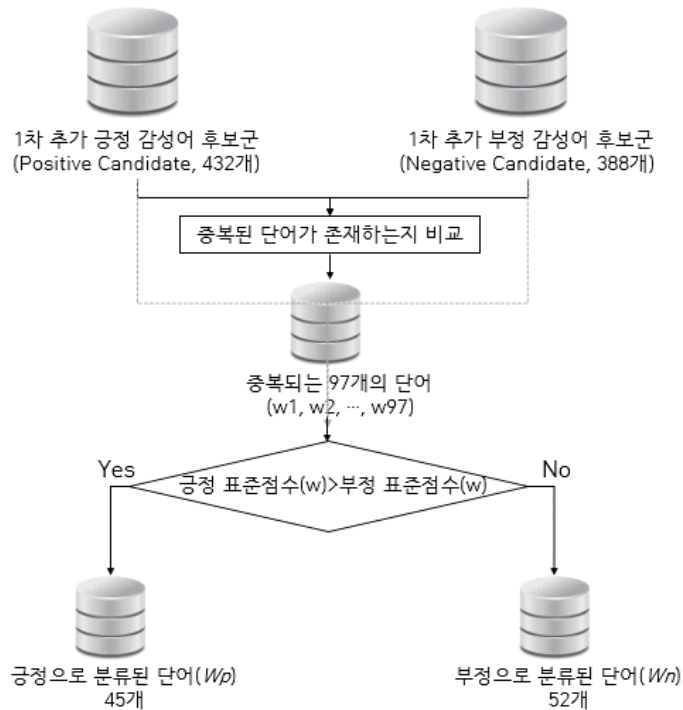


그림 2-7. 1차 추가 긍정 및 부정 감성어 후보군의
중복 단어 분류 과정

따라서 중복되는 97개의 단어의 감성을 결정하기 위하여 각 단어에 대해 출현횟수를 바탕으로 1차 추가 긍정 및 부정 감성어 후보군 내에서의 표준점수(z-score)를 각각 구하고, 이를 비교하여 그 값이 더 큰 감성의 방향으로 결정하였다. 예를 들어, ‘필요/NC’라는 단어는 1차 추가 긍정

및 부정 감성어 후보군에 모두 존재하는 중복단어로 어느 감성에 속하는지 결정해주어야 한다. 따라서 1차 추가 긍정 및 부정 감성어 후보군 내에서의 ‘필요/NC’ 단어의 출현횟수를 바탕으로 표준점수를 각각 구하였다. 그 결과, 긍정의 감성어 후보군 내에서의 표준점수는 -0.39279, 부정의 감성어 후보군 내에서의 표준점수는 2.17124로 나타났다. 따라서 ‘필요/NC’라는 단어는 구글맵 장소 리뷰 텍스트 내에서 긍정적 의미보다는 부정적 의미로 사용된다는 것을 알 수 있으므로, 부정의 감성어로 결정하였다. 이와 같은 방법으로 중복된 97개의 단어에 대해 감성의 방향을 결정한 결과, 긍정으로 분류된 단어는 45개, 부정으로 분류된 단어는 52개로 나타났다.

최종적으로 1차 추가 긍정 감성어 후보군 단어 432개에서 중복 단어 중 부정으로 분류된 단어 52개를 제외하여 총 380개의 추가 긍정 감성어 DB를 생성하였다. 마찬가지로, 1차 추가 부정 감성어 후보군 단어 388개에서 중복 단어 중 긍정으로 분류된 단어 45개를 제외하여 총 343개의 추가 부정 감성어 DB를 생성하였다. 생성된 추가 감성어 DB에는 ‘제주도’, ‘백운계곡’, ‘올림픽공원’과 같은 고유명사가 존재하였는데 이는 구글맵 장소 리뷰 데이터를 수집할 때의 시점에서 해당 장소에 대한 긍정적 또는 부정적 감성이 누적되어 있기 때문이며, 데이터를 수집하는 시점에 따라 추가 감성어 DB에 저장되는 어휘가 달라질 수 있다. 수집한 구글맵 장소 리뷰 가운데 ‘도로에서 백운계곡 쪽으로 진입하기가 어렵습니다’와 같은 특정 장소명을 포함한 부정적 리뷰들이 존재하여 특정 장소명들이 감성어로 추출된 것이다. 이러한 장소명은 감성을 지닌 어휘가 아니며, 이모티콘과 함께 사용되었을 때 의미를 갖는 추가 감성어 DB로 적절하지 않으므로 제거하였다. 따라서 369개의 최종 추가 긍정 감성어 DB와 336개의 최종 추가 부정 감성어 DB를 생성하였다. 표 2-10과 표 2-11은 추가 감성어 DB 일부와 해당 테이블 명세서이며, 추가 감성어 DB의 전체 목록은 <부록 B>에 수록하였다.

본 연구에서 생성한 추가 감성어 DB는 몇 가지 한계가 존재하는데, 구글맵 장소 리뷰 텍스트를 살펴보면 띄어쓰기를 잘 하지 않거나 오타자

가 포함되어 있는 경우가 발견된다. 이는 형태소 분석의 성능을 떨어뜨리는 요인으로 작용하여 ‘잘되어있/NC’, ‘정말별/NC’, ‘없구/NC’, ‘못먹을/NC’, ‘별도많고/NC’, ‘절대가지마/NC’, ‘볼게없고/NC’와 같은 결과가 추가 감성어 DB에 저장되었다. 또한, ‘짬/NC’, ‘공/NC’, ‘안/NC’, ‘어/NC’와 같이 감성을 내포하고 있지 않지만 한 글자로 이루어진 결과도 추가 감성어 DB에 저장되었다. 이는 추가 감성어 DB를 생성하는 과정에서 6자 이상의 단어만을 제외하여 1음절인 결과들이 저장된 것이다. 또한, 형태소 분석을 위해 사용한 R의 KoNLP패키지에서 지원하는 형태소 분석 성능으로 인해 형태소가 부정확하게 태그된 경우가 있다. ‘좋습니/NC’, ‘위치한/NC’, ‘구경하/NC’, ‘블로그에서/NC’, ‘서울에서/NC’가 그 예이다. 이는 추가 감성어 DB의 한계로, 향후 연구를 통해 개선해나가야 한다.

또한, 추가 감성어 DB를 이용하여 더욱 정교한 감성분석을 수행하기 위해서는 추가 감성어 DB에 저장된 각 감성어의 극성뿐 아니라 감성확률을 산출함으로써 이모티콘과 함께 사용된 감성어에 따른 감성의 극성 정도를 고려할 수 있으며 본 연구에서는 이를 향후 연구로 남겨둔다.

표 2-10. 추가 감성어 DB (일부)

추가 긍정 감성어 DB			추가 부정 감성어 DB		
Table Name			Table Name		
ID	SW	Pt	ID	SW	Pt
P_1	많/PA	+1	N_1	없/PA	-1
P_2	공원/NC		N_2	돈/NC	
P_3	아이들/NC		N_3	관리/NC	
P_4	갈/PA		N_4	필요/NC	
P_5	산책하기/NC		N_5	안되/PA	
P_6	놀/PV		N_6	아프/PA	
P_7	가족/NC		N_7	서비스/NC	
P_8	장소/NC		N_8	입장료/NC	
P_9	다양한/NC		N_9	기다리/PV	
P_10	맛있/PA		N_10	망/NC	
...

표 2-11. 추가 감성어 DB 테이블 명세서

Table Name	추가 긍정 감성어 DB, 추가 부정 감성어 DB				
Column Name	Data Type	Nullable	PK	FK	Description
ID	CHAR(10)	NOT NULL	PK		일련번호
SW	VARCHAR(50)	NOT NULL			감성어(sentiment word) 어근(root)과 품사(NC: 명사, PA: 형용사, PV: 동사)가 결합된 형태
Pt	INT	NOT NULL			감성의 극성(polarity) (긍정: 1, 부정: -1)

이모티콘과 추가 감성어 DB를 이용하여 트위터 텍스트에 대한 감성분석을 하기 위해서는 표 2-8에서 정의한 긍정 및 부정 이모티콘과 공간 감성어 사전 및 추가 감성어 DB에 저장된 감성어가 함께 사용이 되었는지를 판단해야 한다. 예를 들어, 표 2-12와 같이 “올림픽공원에서 오랜만에 친구랑 산책하기^^”라는 트위터 텍스트가 있다고 했을 때, 추가 감성어 DB에 저장된 긍정 감성어 ‘산책하기’와 긍정의 이모티콘 ‘^^’이 함께 사용되었으므로, 긍정(+1)의 의미를 갖는다.

표 2-12. 이모티콘과 추가 감성어 DB를 이용한 감성분석 예시

올림픽공원에서 오랜만에 친구랑 산책하기^^	
공간 감성어 사전 및 추가 감성어 DB에 저장된 감성어 산책하기	이모티콘 ^^

이처럼 본 연구에서는 긍정 또는 부정의 감성어와 이모티콘이 함께 사용되는 4가지 경우에 따라 표 2-13과 같이 그 의미를 가진다고 정의하였다. 그림 2-8은 감성어와 이모티콘이 함께 사용된 트위터 텍스트 56개를 바탕으로 긍정 또는 부정의 감성어(w)와 이모티콘(e)이 함께 사용된 분포 결과를 나타낸 것이다. 긍정의 감성어와 긍정의 이모티콘이 함께 사용된 경우가 68%로 가장 많았고, 부정의 감성어와 긍정의 이모티콘이 함께 사용된 경우가 5%로 가장 적었다.

표 2-13. 감성어와 이모티콘이 함께 사용되는 4가지 경우

경우	감성어 w	+	이모티콘 e	=	의미
1	긍정 감성어	+	긍정 이모티콘	=	긍정
예시	산책하기/NC 새롭/PA 운동/NC		^^ ㅎㅎ ~		
2	긍정 감성어	+	부정 이모티콘	=	긍정
예시	좋/PA 추천/NC 그럴듯하/PA		ㅠㅠ ㄷㄷ ...		
3	부정 감성어	+	긍정 이모티콘	=	부정
예시	망/NC 심심/NC 더럽/PA		ㅋㅋ ^^ ♡		
4	부정 감성어	+	부정 이모티콘	=	부정
예시	기다리/PV 맛없/PA 없/PA		:(-- ---		

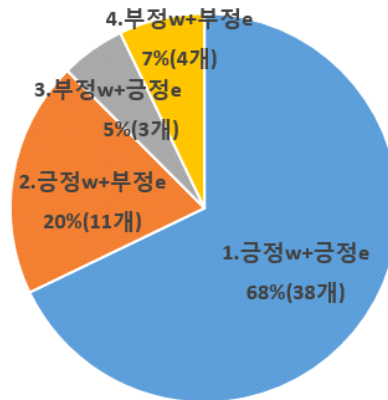


그림 2-8. 트위터 텍스트에서 발견된 감성어와 이모티콘 사용의 분포

이를 바탕으로 긍정 또는 부정의 감성어와 이모티콘이 함께 사용되었을 때, 긍정, 중립, 부정 중 어떤 의미를 갖는지 분류한 결과는 그림 2-9와 같다. 먼저, 긍정의 감성어와 긍정의 이모티콘이 함께 사용되었을 경우, 97%가 긍정의 의미를 나타냈고 나머지 3%는 중립의 의미를 나타냈

다. 둘째로, 긍정의 감성어와 부정의 이모티콘이 함께 사용되었을 경우, 64%가 긍정의 의미를 갖고 나머지 36%가 중립의 의미를 나타냈다. 셋째로, 부정의 감성어와 긍정의 이모티콘이 함께 사용되었을 경우, 67%가 부정의 의미를 나타냈고 33%가 긍정의 의미를 나타냈다. 마지막으로, 부정의 감성어와 부정의 이모티콘이 함께 사용되었을 경우, 75%가 부정의 의미를 갖고 25%가 중립의 의미를 나타냈다. 따라서 본 연구에서는 긍정 또는 부정의 감성어와 이모티콘의 사용 조합에 따라 긍정, 중립, 부정 가운데 50%이상의 의미를 갖는 4가지 경우에 대하여 고려하였다(표 2-13).

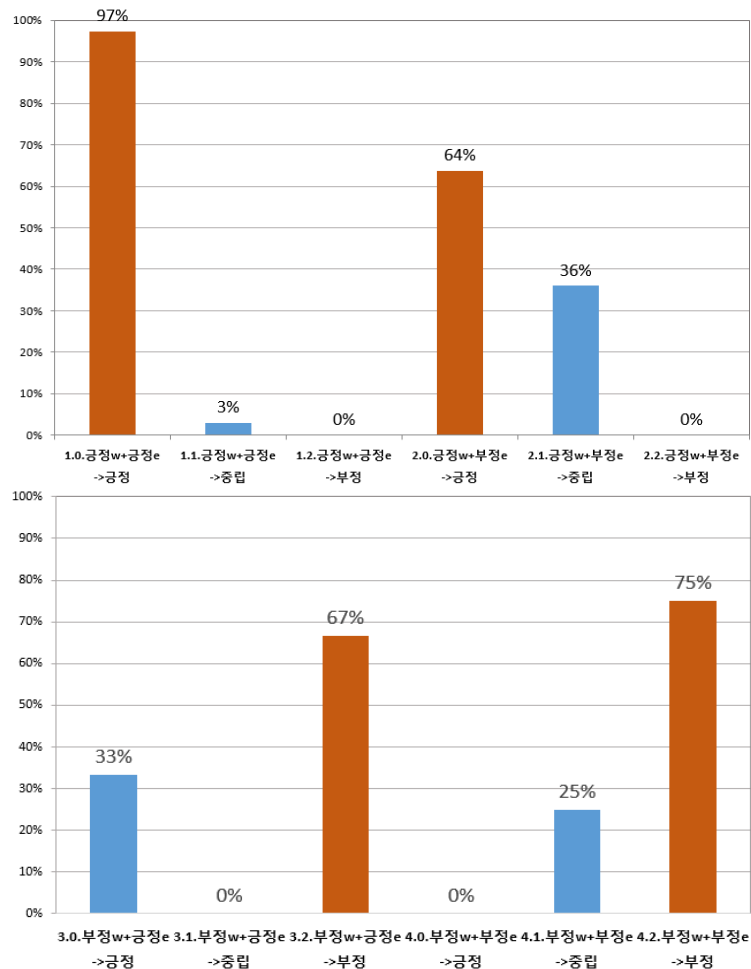


그림 2-9. 감성어와 이모티콘의 조합에 따른 의미 분류 결과

이러한 방식으로 이모티콘과 추가 감성어 DB를 이용한 감성분석을 적용하기 위해서 표 2-6에서 정의한 $PT(i)$ 와 $NT(i)$ 수식에 이모티콘에 의한 감성값을 추가하여 새롭게 정의하면 표 2-14와 같다.

표 2-14. 감성지수벡터 및 감성값 정의

	장소 POI에 대한 감성지수벡터(\vec{S})	텍스트의 감성값(T)
긍정 (Positive)	$\vec{S}_P = \sum_{i=1}^{F_k} PT_{(i)}$	$PT_{(i)} = \sum_{j=1}^n \{match(i,j) \times Pb_{(j)}\} + \sum_{e=1}^m match(i,e,w)$
부정 (Negative)	$\vec{S}_N = \sum_{i=1}^{F_k} NT_{(i)}$	$NT_{(i)} = - \sum_{j=1}^n \{match(i,j) \times Pb_{(j)}\} - \sum_{e=1}^m match(i,e,w)$

표 2-14의 $PT(i)$ 와 $NT(i)$ 를 계산하는 수식에서 새롭게 더해진 $\sum_{e=1}^m match(i,e,w)$ 는 이모티콘을 고려한 감성분석 결과를 적용한 것이다.

여기서, i 는 트위터 텍스트, e 는 표 2-8에서 정의한 긍정 및 부정 이모티콘, m 은 i 에 존재하는 e 의 개수, w 는 공간 감성어 사전 및 추가 감성어 DB에 저장된 감성어를 의미한다. $match(i,e,w)$ 는 아래와 같이 경우에 따라 0 또는 1의 값을 가진다.

$$match(i,j) = \begin{cases} 1 : \text{텍스트 } i \text{에 포함된 } e \text{가 } w \text{와 함께 사용되었을 경우} \\ 0 : \text{그 외의 경우} \end{cases}$$

표 2-15. 트위터 텍스트 감성분석 예시

Text (i)	$PT(i)$ (Positive index of Text(i))	$NT(i)$ (Negative index of Text(i))
날씨가 더워서 [힘드네 $\pi\pi$] 그래도 [아이들 데리고 왔더니 [좋아]하네^] (부정(-1), 1.0) (긍정(+1), 1.0) (긍정(+1), 1.0)	$1+1 = 2$	-1

표 2-15와 같이 트위터 텍스트 예시가 있다고 했을 때, 공간 감성어 사전에 의해 [아이들/NC, 좋/PA]의 조합이 발견되었으므로 PT에 $+1(+1 \times 1.0)$ 을 부여하고, 이모티콘에 의해 [힘들/PV, $\pi\pi$]와 [좋/PA, ^]의 조합이 발견되었으므로 각각 NT에 $-1(-1 \times 1.0)$ 과 PT에 $+1(+1 \times 1.0)$ 을 부여한다. 이 때, 이모티콘의 감성확률(극성의 정도)은 모두 1.0을 부여한다. 즉, 긍정의 이모티콘은 100% 긍정의 의미를, 부정의 이모티콘은 100% 부정의 의미를 지닌다고 정의한다.

따라서 해당 텍스트에서 긍정적 감성정보는 2개([아이들/NC, 좋/PA], [좋/PA, ^])가 발견되었으며, 두 감성값의 합인 2는 PT가 된다. 또한, 부정적 감성정보는 1개([힘들/PV, $\pi\pi$])가 발견되었으며 해당 감성값인 -1이 NT가 된다. 이렇게 모든 트위터 텍스트에 대하여 감성분석을 수행하여 PT와 NT를 산출하고, 이를 각 장소 POI에 해당하는 트위터 텍스트들에 대하여 각각 합산한 값을 긍정지수벡터와 부정지수벡터로 부여한다.

표 2-16은 ‘올림픽공원’에 대한 이모티콘이 포함된 트위터 텍스트의 감성분석 예시이다. 각 텍스트에 대해 이모티콘까지 고려하여 감성분석을 한 결과, $\overrightarrow{S_P} = +11.87$, $\overrightarrow{S_N} = -2.89$ 의 결과를 보인다. 이는 공간 감성어 사전만 이용하여 감성분석한 표 2-7의 결과($\overrightarrow{S_P} = +6.87$, $\overrightarrow{S_N} = -1.89$)와 비교했을 때, 이모티콘을 고려함으로써 텍스트 내의 긍정 및 부정의 감성을 조금 더 세밀하게 포착하였음을 알 수 있다.

표 2-16. ‘올림픽공원’에 대한 감성분석 예시(이모티콘 고려)

i (번호)	Text (i)	감성어 추출 및 개별 텍스트의 감성값 산출	
		$PT(i)$ (Positive index of Text(i))	$NT(i)$ (Negative index of Text(i))
1	급 번개 [나들이♥ 올림픽공원] (긍정(+1), 1.0) (긍정(+1), 1.0)	1.0+1.0 = 2	NA
2	올림픽공원 수영장 다녀오다가 하늘이 맑아 찍어본 하늘사진	NA	NA
3	[주차장도 넓]고 [주차비용도 싸]서 [데이트하러 오기 좋다~] (긍정(+1), 0.98) (긍정(+1), 0.89) (긍정(+1), 1.0) (긍정(+1), 1.0)	0.98+0.89+ 1.0+1.0 = 3.87	NA
4	오늘은 회사언니랑 [올림픽공원까지 산책하러 :)] (긍정(+1), 1.0) (긍정(+1), 1.0)	1.0+1.0 = 2	NA
5	[사람 많]고 너무 [힘들었던 하루 올림픽공원] (중립(0), 0.685) (부정(-1), 1.0)	NA	-1.0
6	올림픽공원 행사하는지 [차가 너무 많다ㅠㅠ].그래도 [단풍은 예쁘네 ㅋㅋㅋ] (부정(-1), 0.89) (부정(-1), 1.0) (긍정(+1), 1.0) (긍정(+1), 1.0)	1.0+1.0 = 2	-0.89-1.0 = -1.89
7	[아이들 데리고 왔더니 좋아하네 ㅎㅎ] 자주 와야지 올림픽공원 (긍정(+1), 1.0) (긍정(+1), 1.0)	1.0+1.0 = 2	NA
<p>빨강: 부정 초록: 중립 파랑: 긍정</p> $\sum_{i=1}^{F_k} PT(i) \text{ or } NT(i)$		+11.87 $\overrightarrow{(S_P)}$	-2.89 $\overrightarrow{(S_N)}$

2.4 장소 선호도 산출

2.4.1 장소 선호도 산출식

본 연구에서는 각 장소 POI에 대한 언급빈도수와 감성분석 결과를 모두 고려하여 장소 선호도를 계량화하기 위하여 장소 선호도 산출식을 제안한다. 식 2-1은 장소 POI k 에 대한 1차 선호도를 산출하는 식이다. 이는 특정 장소 POI에 대해 SNS 사용자들이 많이 방문하고 관심이 높을수록 언급빈도수가 높으며, 장소에 대한 긍정적인 의견 및 감성이 많이 나타날수록 긍정의 감성지수가 높게 나타나므로 해당 장소 POI의 선호도는 높게 나타나는 개념이다. 1차 선호도 산출식을 제안함에 있어서 각 장소 POI k 의 언급빈도수(F_k)에 따라 최대 2배의 점수를 가질 수 있도록 언급빈도수에 대한 가중치 $Weight(F_k)$ 를 설정한다. 이는 명채석(2008)의 연구에서 제안한 상품평 리뷰 개수에 따른 가중치 산정식을 따른다.

$$1차선호도(k) = 감성방향(SO) \times 감성점수(SS) \times Weight(F_k) \quad (2-1)$$

여기서,

$$감성방향(SO, Sentiment Orientation) = \begin{cases} +1, & \|\vec{S_p}\| \geq \|\vec{S_N}\| \\ -1, & \|\vec{S_p}\| < \|\vec{S_N}\| \end{cases} \quad (2-2)$$

즉, 감성방향은 장소 POI k 의 긍정지수벡터($\vec{S_p}$)와 부정지수벡터($\vec{S_N}$)의 크기비교를 통해 긍정(+1) 또는 부정(-1)으로 결정된다.

$$감성점수(SS, Sentiment Score) = 1 + \|\vec{S_p} + \vec{S_N}\| \quad (2-3)$$

감성점수는 장소 POI k 의 긍정지수벡터와 부정지수벡터의 합벡터의 크기에 1을 더한 값이다. \vec{S}_P 와 \vec{S}_N 의 크기가 동일한 중립의 경우, $\|\vec{S}_P + \vec{S}_N\|$ 가 0이 되므로 언급빈도수가 존재하더라도 1차 선호도가 0이 될 수 있으므로, 이에 1을 더하여 언급빈도수가 존재하는 장소 POI k 의 1차 선호도가 0이 되지 않도록 한다.

$$Weight(F_k) = 2 - 2^{-(F_k/wf)} \quad (2-4)$$

언급빈도수에 따른 가중치 $Weight(F_k)$ 는 명재석(2008)의 연구에서 제안한 상품평 리뷰 개수에 따른 가중치를 따르며, 이는 많은 상품평을 갖고 있는 상품에 더 높은 점수를 부여해야 할 필요가 있다고 보아 상품평의 개수가 많을 경우 2에 가까운 값을, 적을 경우 1에 가까운 값을 갖는다. 이때, 가중치 값의 증가 속도를 조절하는 가중치 조절 계수(wf)를 사용했으며, 50으로 그 값을 정한다. 따라서 식 2-4에서 각 장소 POI k 의 언급빈도수(F_k)에 따른 가중치 $Weight(F_k)$ 는 1에서 2사이의 값을 갖는다.

식 2-1을 통해 산출한 장소 POI k 에 대한 1차 선호도(k)는 값의 범위에 따라 해석이 달라지며, 1차 선호도(k)값이 0보다 클 경우 해당 장소 POI k 는 긍정, 0일 경우 중립, 0보다 작을 경우 부정을 의미한다.

이에 따라 표 2-16의 ‘올림픽공원’에 대한 예시에 따른 올림픽공원의 1차 선호도를 구하면 표 2-17과 같다. 올림픽공원의 1차 선호도는 10.90으로 0보다 크므로 올림픽공원은 긍정적인 장소 POI로 해석할 수 있다.

표 2-17. ‘올림픽공원’에 대한 예시에 따른 1차 선호도 산출

감성방향(SO)	감성점수(SS)	$Weight(F_k)$	1차 선호도 (올림픽공원)
+1(긍정)	$1 + 11.87 - 2.89 = 9.98$	$2 - 2^{-(7/50)} = 1.092$	$+1 \times 9.98 \times 1.092 = 10.90$

2.4.2 장소 유형별 선호도 정규화

선호도를 분석하고자 하는 장소 유형, 예를 들어, 공원, 카페, 쇼핑몰, 관광지 가운데 한 가지를 선택하여 해당 장소 유형에 속하는 각 장소 POI의 1차 선호도를 분석할 수 있다. 최종적으로 해당 장소 유형에 속하는 여러 장소 POI 가운데 1차 선호도가 가장 높은 POI부터 1차 선호도가 가장 낮은 POI까지 도출된다. 장소 유형에 따라 해당 장소 POI들의 1차 선호도 값의 범위가 다르기 때문에 이들 간의 비교를 위해서 정규화(min-max normalization)할 필요가 있다. 따라서 1차 선호도를 식 2-5에 의해 정규화함으로써 최종 장소 선호도(P)를 도출할 수 있다.

$$P(k) = \frac{1차선호도(k) - 1차선호도_{\min}}{1차선호도_{\max} - 1차선호도_{\min}} \times 100 \quad (2-5)$$

정규화를 통해 0~1 사이의 값이 도출되며, 하위의 선호도 값은 0.0091과 같이 소수점 셋째자리부터 0이 아닌 숫자가 나타나는 작은 값일 경우가 있으므로 이에 100을 곱하여 값의 크기를 증가시켰다. 이에 따라 장소 POI k 의 최종 선호도 $P(k)$ 는 0부터 100사이의 값을 갖는다. 최종 선호도 $P(k)$ 가 100에 가까울수록 해당 장소 유형에 속하는 여러 장소 POI 가운데 선호도가 높다는 것을 의미한다.

3. 실험 적용 및 결과

3.1 실험 대상 및 데이터

본 연구에서 제안한 장소 선호도 분석기법을 통해 실제 트위터 텍스트를 대상으로 서울에 위치한 ‘테마공원’ 105곳에 대해 선호도 분석을 수행하였다. 이를 위해서 먼저, 네이버 검색 API를 이용하여 ‘서울 테마공원’으로 검색했을 때 도출되는 테마공원 POI 138개를 수집하였다. 이 가운데 테마공원이라고 보기 어려운 ‘어린이공원 관리사무소, 하늘공원 사주카페, 테마노래 연습장, 홀리랜드익스피어리언스, 홍토피아엔, 동일스위트리버 아파트테마공원, 목동현대하이퍼리온 2차아파트테마공원, 펫펫골프앤게임, 해피썸머페스티벌 시즌4키즈파크, 한국자연공원협회’ 등 33개의 POI를 제거하였다. 제거 후의 105개의 서울 테마공원 POI는 표 3-1과 같다.

실험에 이용한 트위터 데이터는 트위터 오픈 API를 이용하여 2015년 7월 22일부터 2016년 2월 26일까지 약 7개월 간 수집한 트위터 데이터 273,515개이다.

다음 절에서는 수집한 트위터 데이터를 이용하여 105개의 서울 테마공원 POI의 선호도를 분석하는 과정에 대해 설명한다.

표 3-1. 서울 테마공원 POI

번호	POI	번호	POI	번호	POI
1	도산공원	41	오동공원	81	황금내공원
2	월드컵공원	42	뽕고개공원	82	양화폭포
3	서대문독립공원	43	난지천공원	83	박시환숲
4	허브천문공원	44	갈산공원	84	세종문화회관 예술의정원
5	효창공원	45	사기정공원	85	구암공원
6	파리공원	46	수락산당고개지구공원	86	손기정공원
7	서서울호수공원	47	등촌공원	87	금천체육공원
8	탑골공원	48	성수공원	88	동대문역사관 1398
9	노을공원	49	비석거리공원	89	올림픽공원
10	남산공원	50	현대원서공원	90	광나루자전거공원
11	와룡공원	51	북악팔각정	91	롯데월드
12	동대문역사문화공원	52	일자산자연공원	92	어린이대공원
13	몽마르뜨공원	53	삼선공원	93	성수동구두테마공원
14	남산골 전통정원	54	중마루공원	94	서남물재생센터공원 테마숲
15	학동공원	55	염창공원	95	용마랜드
16	오목공원	56	누에머리공원	96	서울색공원
17	천호공원	57	세종로공원	97	자연순환테마전시관
18	낙성대공원	58	온수공원	98	우이공원유원지
19	장충단공원	59	각심어린이공원	99	용마공원놀이동산
20	서초문화예술공원	60	까치공원	100	한강랜드
21	용마폭포공원	61	송이공원	101	동막골유원지
22	양카라공원	62	대모산자연공원	102	광나루유원지
23	대한산배수지공원	63	불암허브공원	103	우이동유원지
24	영등포공원	64	보람어린이공원	104	아차산성유원지
25	서울창포원	65	정동공원	105	도봉유원지
26	오금공원	66	경희궁공원		
27	역삼공원	67	매화공원		
28	발바닥공원	68	백연공원		
29	우장산근린공원	69	서낭당공원		
30	북서울꿈의숲	70	어린이교통공원		
31	양재근린공원	71	연의공원		
32	훈원원공원	72	동작주차공원		
33	살곶이체육공원	73	방화근린공원		
34	관악산호수공원	74	노량진근린공원		
35	삼일공원	75	홍제근린공원		
36	간테메공원	76	장안어린이공원		
37	장미공원	77	솔피공원		
38	동명근린공원	78	여의도 한강공원		
39	방화공원	79	호현당		
40	평화의공원	80	청계천베를린광장		

3.2 서울 테마공원 선호도 분석

3.2.1 서울 테마공원 POI에 대한 언급빈도수 산출

앞서 수집한 트위터 데이터 가운데 사용자가 남긴 트위터 텍스트 정보가 담긴 'text' 필드를 리스트 형태로 저장하여 273,515개의 트위터 텍스트 말뭉치를 생성하였다. 다음으로, 텍스트 내에 포함된 'RT, #, @'과 URL 주소 그리고 숫자 및 영문자와 같은 불용어를 제거하였다. 불용어가 제거된 트위터 텍스트 가운데 앞서 생성한 105개의 서울 테마공원 POI(표 3-1)를 포함하는 트위터 텍스트를 추출하였다. 이를 위해 키워드 매칭 기능을 제공하는 R의 stringr 패키지를 이용하였다.

그 결과, 1,710개의 트위터 텍스트가 추출되었으며, 트위터 텍스트 내에서 한 번이라도 언급된 서울 테마공원 POI의 개수는 105개 중 36개로 이를 대상으로 선호도를 분석하였다(표 3-2). 이 가운데 '여의도한강공원'의 경우, 트위터 텍스트에서 발견되는 동일한 장소를 지칭하는 '한강 여의도 공원', '여의도한강시민공원', '여의나루 한강공원'의 명칭을 모두 고려하여 '여의도한강공원'의 언급빈도수를 산출하였다.

서울 테마공원 POI를 포함하고 있는 트위터 텍스트의 분석을 위해서는 자연어 처리 과정을 거쳐야 한다. 따라서 형태소 분석을 수행하여 텍스트를 분해 가능한 최소한의 단위로 분리하여 품사를 부여하였다. 이를 위해 R에서 한글 자연어 분석을 지원하는 KoNLP(Korean Natural Language Processing)패키지를 이용하였다. 한나눔 형태소 분석기에서 사용하는 카이스트 형태소 태그 집합⁷⁾을 바탕으로 각 단어에 대해 22 종류의 품사를 부여하는 KoNLP 패키지의 SimplePos22 함수를 이용하여 각 트위터 텍스트에 대한 형태소 분석을 수행하였다.

형태소 분석이 이루어진 트위터 텍스트를 대상으로 각 서울 테마공원 POI의 언급빈도수 즉, 전체 트위터 텍스트 문서 내에 나타나는 각 서울

7) <https://github.com/haven-jeon/KoNLP/blob/master/etc/KoNLP-API.md>
(접속일: 2017년 6월 10일)

테마공원 POI의 총 빈도수를 산출하였다. 그 결과는 표 3-2와 같다.

가장 많이 언급된 테마공원 POI는 롯데월드(750), 올림픽공원(328), 여의도한강공원(137), 동대문역사문화공원(124), 어린이대공원(109) 순으로 나타났으며, 6개의 테마공원 POI는 단 한번만 언급되었다. 이는 앞서 수집한 트위터 데이터의 개수가 풍부하지 않아 언급빈도수가 1,2개만 존재하는 경우가 발생한 것으로, 이는 추후 산출되는 장소 선호도의 정확도를 떨어뜨리는 요인으로 작용할 수 있다. 따라서 향후 오랜 기간 수집한 풍부한 트위터 데이터를 이용하여 각 테마공원 POI가 유의미한 언급빈도수를 가질 수 있도록 해야 할 필요가 있다.

표 3-2. 트위터 텍스트에서 언급된 서울 테마공원 POI 및 해당 언급빈도수

번호	POI	언급빈도수	번호	POI	언급빈도수
1	롯데월드	750	19	구암공원	7
2	올림픽공원	328	20	오목공원	6
3	여의도한강공원	137	21	낙성대공원	5
4	동대문역사문화공원	124	22	영등포공원	5
5	어린이대공원	109	23	서울색공원	4
6	도산공원	49	24	탑골공원	4
7	남산공원	33	25	난지천공원	3
8	월드컵공원	21	26	세종로공원	3
9	북서울꿈의숲	14	27	허브천문공원	2
10	방화근린공원	14	28	북악팔각정	2
11	용마랜드	14	29	장충단공원	2
12	와룡공원	12	30	용마폭포공원	2
13	서서울호수공원	10	31	광나루자전거공원	1
14	효창공원	9	32	몽마르뜨공원	1
15	파리공원	9	33	서울창포원	1
16	노을공원	9	34	천호공원	1
17	학동공원	9	35	평화의공원	1
18	장미공원	8	36	한강랜드	1

3.2.2 서울 테마공원 POI에 대한 감성분석

서울 테마공원 POI의 선호도를 분석하기 위해서 각 테마공원 POI의 언급빈도수를 산출한 후, 감성분석을 수행하였다. 먼저, 앞서 형태소 분석이 이루어진 서울 테마공원 POI를 포함하는 1,710개의 트위터 텍스트와 공간 감성어 사전에 저장된 감성어와의 키워드 매칭을 통해 텍스트 내 감성정보를 추출하였다. 그 과정에서 문맥을 고려하기 위하여 어휘 수식 거리를 적용하여, 매칭되는 감성어를 기준으로 ± 3 어절(word phrase) 거리 내에 테마공원 POI, 지시대명사인 ‘여기’, ‘이곳’ 또는 장소의 속성을 나타내는 공간 주제어가 위치하였을 경우에 해당 감성어를 추출하였다.

또한, 트위터 텍스트의 특성을 반영하기 위해 이모티콘을 중요한 감성정보로 고려하였다. 서울 테마공원 POI를 포함하는 1,710개의 트위터 텍스트 가운데 표 2-8에서 정의한 긍정 및 부정의 이모티콘을 포함하여 작성된 텍스트는 283개로, 약 16.5%의 사용자가 이모티콘을 사용하여 트위터 텍스트를 작성했음을 알 수 있다.

트위터 텍스트 내에서 이모티콘이 어떠한 의미로 사용되었는지 판단하기 위하여 공간 감성어 사전과 앞서 표 2-10에서 정의한 추가 감성어 DB에 저장된 감성어와 이모티콘이 함께 사용되었을 경우 이를 추출하였다.

서울 테마공원 POI를 포함하는 1,710개의 트위터 텍스트로부터 공간 감성어 사전, 추가 감성어 DB와 이모티콘을 고려하여 감성정보를 추출한 결과, 115개의 트위터 텍스트에서 제안한 방법을 통해 감성정보가 추출되었고, 그 외의 트위터 텍스트에서는 감성정보가 발견되지 않았다. 이를 바탕으로 앞서 공간 감성어 사전에 저장된 감성어와의 매칭 과정에서 적용한 3 어절의 어휘 수식 거리를 통해 장소에 대한 표현을 적절하게 포착하였는지 분류한 결과, 75%의 매칭률로 장소에 대한 감성정보를 추출하였음을 알 수 있었다.

사람의 감성은 복잡하고 그 표현 형태도 다양하기 때문에, 장소에 대한 감성인지 사용자의 기분이나 상황에 대한 감성인지 구분하기가 모호

하거나 혼동되는 경우가 있다. 예를 들어, 어떤 사용자가 올림픽공원 체조경기장에서 진행되는 아이돌 그룹 엑소의 콘서트에 참가하여 ‘오빠들 짱!! 올림픽공원 체조경기장’이라는 텍스트를 남긴 경우에, 본 연구에서는 ‘짱’이라는 긍정의 감성정보가 발견되었으므로 긍정의 감성으로 고려하였다. 본 연구에서 제안하는 어휘 수식 거리를 이용한 기계적인 추출 방법을 통해 텍스트로부터 장소에 대한 감성표현만을 정확하게 추출하지 못하는 것은 본 연구가 가지는 한계이자 추후 구문에 대한 분석과 언어학적 접근을 통해 향상시켜야 할 부분이다.

감성정보가 포함된 115개의 트위터 텍스트를 바탕으로 표 2-14에서 정의한 수식과 같이 각 트위터 텍스트별로 추출된 감성어와 이모티콘의 감성값을 긍정과 부정으로 구분하여 누적 합산한 값을 각각 해당 텍스트의 긍정값과 부정값으로 부여하였다. 표 3-3은 그 결과의 일부이며 전체 결과는 <부록 C>에 수록하였다.

표 3-3. 서울 테마공원 POI를 포함하는 트위터 텍스트의 감성분석 결과(일부)

번호	POI	불용어 제거된 트위터 텍스트	공간 감성어 사전		이모티콘 고려		PT	NT
			주어	감성어	공간 감성어 및 추가 감성어	이모티콘		
1	낙성대공원	도서관도 있고 좋네 (낙성대공원 () 서울특별시.)	낙성대공원	좋/PA	NA	NA	1	NA
2	낙성대공원	시작 ... (낙성대공원 () 서울특별시.)	NA	NA	시작/NC	...	1	NA
3	남산공원	내일은 우리 뽀샵군 데리고 산책해야지~ . 산책 공원 매미소리좋다 강아지 개 남산공원 남산공원	NA	NA	산책/NC	~	1	NA
4	남산공원	남산공원 처음으로 갔다. 길 몰라서 재미있었다 ㅋㅋ...	NA	NA	재미/NC	ㅋㅋ	1	NA
5	노을공원	왜 나는 이런 날씨에만 녹화를 나오는가... 태안에서 눈보라 맞은 이후 정확히 년후 상암동 겨울왕국을 만나다. 노을공원	NA	NA	나오/PV	...	NA	-1

번호	POI	불용어 제거된 트위터 텍스트	공간 감성어 사전		이모티콘 고려		PT	NT
			주어	감성어	공간 감성어 및 추가 감성어	이모티콘		
6	노을공원	흐린 날씨... 돈키우스 트레킹 노을공원 노을공원	NA	NA	날씨/NC	...	NA	-1
7	노을공원	아이 좋아~ 노을공원 캠핑장	노을공원	좋/PA	좋/PA	~	2	NA
8	노을공원	급 번개 나들이 :) (노을공원 서울특별시.)	노을공원	나들이 /NC	나들이 /NC	:)	2	NA
9	도산공원	산책 (도산공원 () 서울특별시)	도산공원	산책 /NC	NA	NA	1	NA
10	도산공원	멜팅샵 현대카드 고메워크 ... 바질 사랑해여 ㅋㅋㅋ트트트 소프트셀크랩 카프레제파스타 장백기 봤다 ㅋㅋㅋ트 도산공원...	NA	NA	보/PV	ㅋㅋㅋ	1	NA
...

표 3-3에서 도출된 각 서울 테마공원 POI에 해당하는 트위터 텍스트들의 긍정값과 부정값을 각각 합산함으로써 각 테마공원 POI의 긍정지수벡터와 부정지수벡터를 구하였다.

그 결과, 36개의 서울 테마공원 POI의 긍정지수벡터와 부정지수벡터를 산출한 감성분석 결과는 표 3-4와 같다. SNS 사용자들은 대체적으로 글을 남길 때 장소에 대한 긍정적 감성을 표현하는 경우가 많으므로, 대부분의 테마공원 POI에서 긍정지수벡터가 부정지수벡터의 크기보다 높은 값을 가지는 것을 확인할 수 있다. 또한, 트위터 텍스트 내에서 언급은 되었지만, 감성정보가 발견되지 않은 테마공원 POI는 18개로, 이들의 긍정지수벡터와 부정지수벡터의 크기는 모두 0이다.

3.2.3 서울 테마공원 POI의 선호도 분석 결과

36개의 서울 테마공원 POI의 언급빈도수와 긍정지수백터와 부정지수백터 결과를 바탕으로 식 2-1에서 제안한 장소 선호도 산출식을 통해 1차 선호도를 산출하였다. 다음으로, 장소 유형에 따라 달라지는 해당 장소 POI들의 1차 선호도 값의 범위를 0~100사이의 값으로 표현하기 위하여 식 2-4에 의해 1차 선호도를 정규화함으로써 최종 선호도를 도출하였다. 그 결과는 표 3-4와 같으며 이의 분포는 그림 3-1과 같다.

이 결과는 약 7개월 간 수집된 트위터 데이터로부터 도출한 서울 테마공원 36곳의 장소 선호도로, 올림픽공원, 롯데월드, 여의도한강공원, 도산공원, 동대문역사문화공원, 어린이대공원, 월드컵공원, 북서울꿈의숲, 효창공원 순으로 선호도가 높게 나타났다. 선호도에 미치는 영향력을 비교해본 결과, 감성점수가 선호도에 미치는 영향이 언급빈도수가 미치는 영향보다 전반적으로 높게 나타났다. 올림픽공원의 경우, 롯데월드보다 언급빈도수는 적지만 감성점수가 가장 높아 테마공원 중 가장 높은 선호도를 갖는다. 상위권에 속하는 테마공원들은 대체적으로 인지도가 높은 공원으로 언급빈도수가 높은 경향을 보인다. 그 가운데 4위의 도산공원과 9위의 효창공원, 11위 허브천문공원은 상위권에 속하는 다른 테마공원들에 비해 언급빈도수는 적지만 감성점수의 영향으로 상위권에 올랐다. 반면, 18위의 방화근린공원부터 36위의 용마폭포공원은 감성정보가 발견되지 않아 언급빈도수에 따라 선호도가 결정되었으며, 최하위의 용마폭포공원은 부정지수백터만 존재하여 가장 선호도가 낮게 나타났다. 이는 풍부하지 못한 트위터 데이터 개수의 한계로 인해 한 자릿수의 언급빈도수로 선호도가 결정되는 결과를 보였다. 따라서 추후 더욱 더 많은 트위터 데이터를 확보하여 선호도를 산출할 필요가 있다.

본 연구의 분석 대상인 테마공원 외의 다른 유형의 장소에 대해 선호도를 분석한다면 다른 결과가 도출될 것이다. 예를 들어, ‘맛집’을 대상으로 선호도를 분석한다면 텍스트에서 발견되는 감성정보는 ‘맛있다’, ‘청결하다’, ‘양이 많다’와 같은 음식점 또는 음식에 대한 표현일 것이다. 또한,

긍정적인 내용이 주로 작성되는 테마공원과 비교했을 때, 맛집에 대한 텍스트에서는 긍정적 표현뿐 아니라 부정적 표현 또한 다수 발견될 것으로 예상된다. 사용자들은 다녀온 음식점에 대하여 ‘대기시간이 길다’, ‘맛이 없어졌다’, ‘가격이 비싸다’와 같은 부정적 평가도 자주 작성하므로 테마공원의 분석결과와 비교했을 때, 긍정지수벡터뿐 아니라 부정지수벡터의 크기 또한 높게 나타나는 POI가 존재할 것이다. 이러한 분석결과는 해당 업체가 개선할 점을 파악하는 데에도 도움이 될 수 있다. 이러한 분석을 가능하게 하기 위해서는 본 연구에서 사용한 공간 감성어 사전이 아닌 ‘음식점’에 대한 감성표현을 저장한 감성어 사전이 필요할 것이다. 즉, 분석대상에 따른 감성어 사전 구축 결과를 바탕으로 분석이 이루어져야 한다.

다음 절에서는 본 연구에서 제안한 장소 선호도 분석기법을 통해 추출된 감성정보를 포함한 트위터 텍스트의 감성분석 수행 결과에 대한 정확도를 평가한다.

표 3-4. 서울 테마공원 POI의 선호도 분석 결과

순위	POI	\vec{S}_P	\vec{S}_N	언급 빈도수 (F_k)	Weight (F_k)	1차 선호도	최종 선호도	긍정 키워드	부정 키워드
1	올림픽공원	24.96	0	328	1.99	51.64	100.00	걷다, 좋다, 짱, 운동, 나들이, 즐겁다, 대박	
2	롯데월드	24.95	-5.99	750	2.00	39.92	78.16	멋지다, 신기하다, 아름답다, 나들이, 최고, 깨끗, 놀다, 천국, 넓다, 좋다, 멋있다	아쉽다, 힘들다
3	여의도 한강공원	13	-2	137	1.85	22.20	45.17	나들이, 좋다, 맑다, 이쁘다	힘들다
4	도산공원	8.96	0	49	1.49	14.87	31.52	조용하다, 산책, 보다, 운동, 대박, 천국	
5	동대문역사 문화공원	4	0	124	1.82	9.10	20.78	예쁘다, 하다	
6	어린이대공원	4	-1	109	1.78	7.12	17.08	좋다, 나들이	안된다
7	월드컵공원	4	0	21	1.25	6.26	15.49	나들이, 좋다, 산책, 바라보다	
8	북서울꿈의숲	3	0	14	1.18	4.71	12.59	숲, 보다	
9	효창공원	2.99	0	9	1.12	4.46	12.13	상쾌, 활동	
10	남산공원	2	0	33	1.37	4.10	11.46	산책, 재밌다	
11	허브천문공원	2.97	0	2	1.03	4.08	11.42	아담하다, 좋다	
12	노을공원	4	-2	9	1.12	3.35	10.07	좋다, 나들이	나오다
13	낙성대공원	2	0	5	1.07	3.20	9.79	좋다	
14	파리공원	1	0	9	1.12	2.23	7.99	예쁘다	
15	오목공원	1	0	6	1.08	2.16	7.85	좋다	
16	영등포공원	1	0	5	1.07	2.13	7.80	지나가다	
17	몽마르트공원	0.98	0	1	1.01	2.01	7.56	천국	
18	방화근린공원	0	0	14	1.18	1.18	6.02		
18	용마랜드	0	0	14	1.18	1.18	6.02		
20	와룡공원	0	0	12	1.15	1.15	5.97		

순위	POI	\vec{S}_P	\vec{S}_N	언급 빈도수 (F_k)	$Weight(F_k)$	1차 선호도	최종 선호도	긍정 키워드	부정 키워드
21	서서울 호수공원	0	0	10	1.13	1.13	5.93		
22	학동공원	1	-1	9	1.12	1.12	5.91	맑다	심심하다
23	장미공원	0	0	8	1.10	1.10	5.88		
24	구암공원	0	0	7	1.09	1.09	5.86		
25	서울색공원	0	0	4	1.05	1.05	5.79		
25	탑골공원	0	0	4	1.05	1.05	5.79		
27	난지천공원	0	0	3	1.04	1.04	5.76		
27	세종로공원	0	0	3	1.04	1.04	5.76		
29	북악팔각정	0	0	2	1.03	1.03	5.74		
29	장충단공원	0	0	2	1.03	1.03	5.74		
31	광나루자전거공원	0	0	1	1.01	1.01	5.71		
31	서울창포원	0	0	1	1.01	1.01	5.71		
31	천호공원	0	0	1	1.01	1.01	5.71		
31	평화의공원	0	0	1	1.01	1.01	5.71		
31	한강랜드	0	0	1	1.01	1.01	5.71		
31	용마폭포공원	0	-1	2	1.03	-2.05	0.00		모르겠다

계급구간	빈도수
5	1
10	23
15	5
20	2
25	1
30	0
35	1
40	0
45	0
50	1
55	0
60	0
65	0
70	0
75	0
80	1
85	0
90	0
95	0
100	1
기타	0

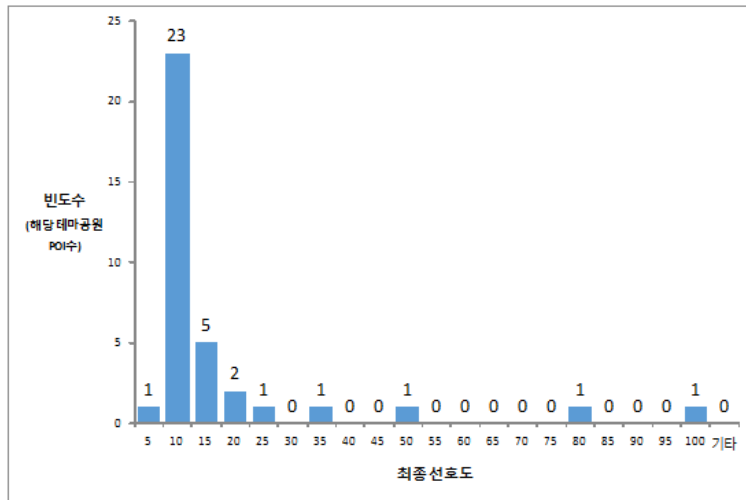


그림 3-1. 서울 테마공원 POI의 선호도 분포

3.3 서울 테마공원 POI에 대한 감성분석 정확도 평가

본 연구에서 제안한 방법을 통해 추출된 감성정보를 포함한 115개의 트위터 텍스트에 대하여 문장의 감성 분류 성능을 평가하기 위해 정확률(precision)과 재현율(recall)을 이용한 F1-measure을 사용하였다. 먼저, 이를 위해 SNS를 활발하게 사용하는 20, 30대를 대상으로 설문조사를 수행하였다. 감성정보를 포함하는 115개의 트위터 텍스트에 대하여 해당 장소에 대한 긍정, 중립, 부정의 감성 중 하나를 선택하도록 하였다. 총 82건의 응답을 수집하였으며, 그 중 1건의 응답은 115개의 문항 중 16개의 문항에만 답을 한 불성실한 응답으로 제외하였다. 그 결과, 81건의 응답을 얻었으며(<부록 D> 참조), 설문에 대한 신뢰도를 평가하기 위하여 SPSS를 이용하여 Cronbach's α (신뢰도 계수, reliability coefficient)⁸⁾를 구하였다. 이를 위해 응답의 긍정을 +1, 중립을 0, 부정을 -1의 값으로 바꾸고 SPSS의 신뢰도 분석 도구를 실행하였다. 문항들이 하나의 개념을 얼마나 잘 표현하는지는 내적 일치도(internal consistency)를 통해 확인할 수 있으며, 내적 일치도를 계산한 값을 Cronbach's α 라 하며, 각 문항을 하나의 검사로 간주하여 신뢰도를 추정하므로 문항의 신뢰도라 하기도 한다(권세혁, 2007). 식 3-1은 Cronbach's α 를 구하기 위한 식이다.

$$\alpha = \frac{K}{K-1} \left(1 - \frac{\sum_{i=1}^K \sigma_{Y_i}^2}{\sigma_X^2} \right) \quad (3-1)$$

여기서, K 는 문항 수, $\sigma_{Y_i}^2$ 는 i 번째 문항 점수의 분산, σ_X^2 는 총점의 분산을 의미한다. Cronbach's α 는 0에서 1 사이의 값을 가지며, 1에 가까울수록 신뢰도가 높다. 문항의 수가 많을수록, 응답자 수가 많을수록 값이 높아지는 경향이 있다. 115개의 항목에 대한 81건의 설문 응답의 신

8) http://zetawiki.com/wiki/%ED%81%AC%EB%A1%A0%EB%B0%94%ED%9D%90_%EC%95%8C%ED%8C%8C#.EA.B0.9C.EC.9A.94 (접속일: 2017년 6월 10일)

되도는 0.962의 Cronbach's α 을 가지며, 이는 본 설문의 신뢰도가 높다고 해석될 수 있다. 81건의 응답 결과를 종합하여 각 트위터 텍스트에 대한 설문 결과(actual)와 본 연구에 따른 분석 결과(predicted)를 비교하였다. 즉, 감성정보를 포함한 115개의 트위터 텍스트를 대상으로 본 연구에 의해 판단된 각 트위터 텍스트의 감성의 극성(긍정, 중립, 부정)이 설문조사에 의한 감성판단 결과와 일치하는지 비교분석하였다. 설문 응답 결과를 바탕으로 각 트위터 텍스트에 대하여 긍정, 중립, 부정 중 가장 많은 선택을 받은 감성을 해당 트위터 텍스트의 감성(actual)으로 결정하였다. 식 3-2, 식 3-3, 식 3-4를 바탕으로 트위터 텍스트 문장의 감성 분류 성능을 평가하기 위해 정확률(precision)과 재현율(recall)을 이용한 F1-measure을 구하였다.

$$\text{정확률}(p) = \frac{\text{해당 감성으로 분류된 실제 해당 감성 텍스트 수}}{\text{해당 감성으로 분류된 텍스트 수}} \quad (3-2)$$

$$\text{재현율}(r) = \frac{\text{해당 감성으로 분류된 실제 해당 감성 텍스트 수}}{\text{해당 감성 전체 텍스트 수}} \quad (3-3)$$

$$F_1 - \text{measure} = 2 \times \frac{p \times r}{p + r} \quad (3-4)$$

그 결과는 표 3-5, 표 3-6과 같다. 115개의 트위터 텍스트 가운데 긍정, 중립, 부정으로 올바르게 감성분석이 수행된 트위터 텍스트는 90개로, 약 78.26%의 정확도(overall accuracy)를 보였다. 이는 감성어 사전을 이용하여 트위터 데이터의 감성분석을 수행한 오평화(2016) 연구의 감성분석 결과 정확도인 71% 그리고 기계학습을 통해 트위터 데이터의 감성분류를 수행한 임좌상(2014) 연구의 정확도인 76%와 비교하였을 때, 향상된 결과를 보인다. 따라서 본 연구에서 제안한 방법에 의해 장소 선호도를 도출하는 것이 의미가 있다고 할 수 있다.

표 3-5. 감성분석 정확도 평가

		트위터 텍스트 감성분석 결과(Predicted)				
		긍정	중립	부정	합계	Recall
설문결과 (Actual)	긍정	80	2	4	86	0.9302
	중립	11	7	6	24	0.2917
	부정	0	2	3	5	0.6000
	합계	91	11	13	115	
	Precision	0.8791	0.6364	0.2308		78.26% (overall accuracy)

표 3-6. 감성별 정확도 평가

감성	정확률(precision)	재현율(recall)	F1-measure
긍정	0.8791	0.9302	0.9040
중립	0.6364	0.2917	0.4000
부정	0.2308	0.6000	0.3333

또한, 트위터 텍스트 문장의 감성 분류 성능을 평가하기 위하여 F1-measure를 살펴보면, 긍정, 중립, 부정 가운데 긍정에 대한 F1-measure가 가장 높고, 중립과 부정에 대한 F1-measure는 낮게 나타났다. 긍정의 경우 정확률과 재현율이 거의 유사하게 산출되었으나, 중립의 경우 정확률이, 부정의 경우 재현율이 더 높게 나타났다. 즉, 중립의 경우 중립으로 추출된 결과 중 중립이 맞는 경우가, 부정의 경우 실제 부정 가운데 부정으로 추출된 경우가 더 많음을 나타낸다. 반대로는, 중립의 경우 실제 중립 중 중립으로 추출하는 성능이 낮고, 부정의 경우 부정으로 추출된 결과 중 실제 부정이 맞을 정확률이 낮음을 의미한다. 이는 감성어 사전에 저장된 한정된 감성어의 한계와 트위터 사용자들이 글을 작성할 때 긍정적인 감성이나 경험을 주로 작성하기 때문인 것으로 보인다. 향후 감성어 사전을 확장하고, 형태소 분석 및 구문 분석의 자연어처리 성능 향상을 통해 정확률과 재현율을 동시에 높이는 연구를 수행할 필요가 있다.

4. 장소 선호도 분석 활용 방안

본 연구에서 제안하는 장소 선호도 분석 기법은 SNS 텍스트를 비롯한 비정형의 텍스트 데이터로부터 장소 선호도 분석을 가능하게 한다. 이는 사용자가 직접 웹사이트에 접속하여 남기는 장소에 대한 별점과 리뷰뿐 아니라 일상에서 작성하는 다양한 포스트, 블로그, 카페글 등의 텍스트 데이터로부터 장소 선호도를 분석할 수 있다는 점에서 의의가 있다. 이에 본 연구에서 제안하는 기법을 통해 도출된 장소 선호도 결과를 현재 장소 정보 제공 서비스에 활용하기 위한 방안을 살펴보았다.

먼저, 네이버 지도의 장소 검색 결과는 그림 4-1 및 그림 4-2와 같다.

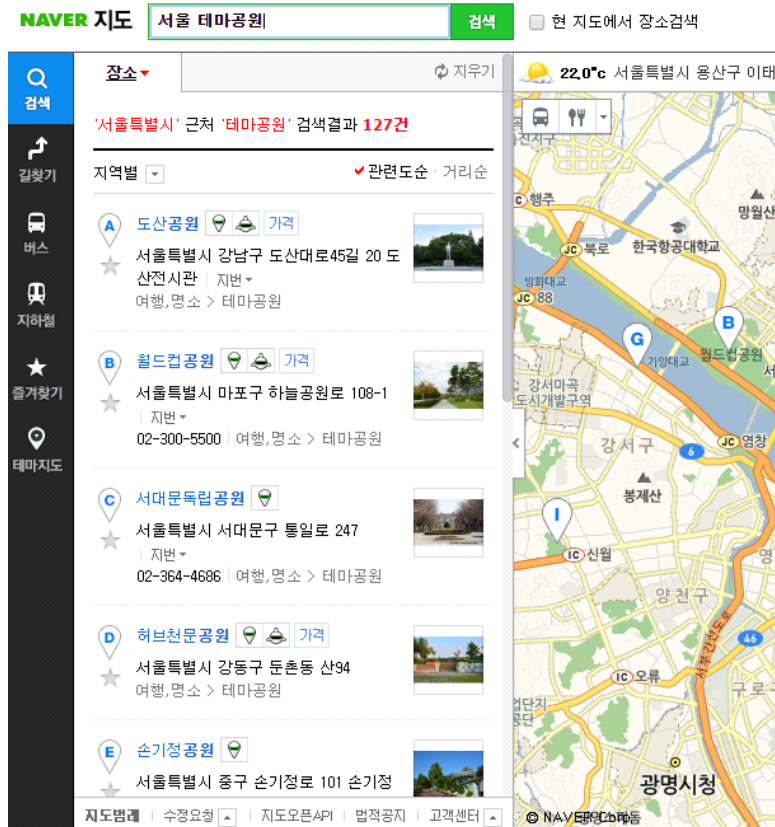


그림 4-1. 네이버 지도의 장소 검색 결과(PC 버전)

네이버 지도 서비스에서는 검색어에 따른 POI 목록을 제공하며, 평점 정보는 따로 제공하고 있지 않다. 모바일 버전 검색 결과의 경우(그림 4-3), 관련도순, 거리순과 더불어 리뷰순의 결과 정렬을 제공하고 있다. 그러나 이는 단순히 각 POI에 대한 네이버 블로그, 포스트 수가 많은 순으로 제공하고 있다. 따라서 본 연구에서 제안하는 장소 선호도 분석기법을 활용하여 리뷰수뿐 아니라 네이버 블로그, 포스트의 텍스트 내용을 기반으로 선호도를 분석하여 장소 선호도가 반영된 장소 검색 결과 제공이 가능할 것이다.

월드컵공원

지도보기 >

길찾기 >

주소

서울특별시 마포구 하늘공원로 108-1
[지번] 서울특별시 마포구 상암동 481-6

전화번호

02-800-5500 더보기

분류

여형,명소 > 테마공원

홈페이지

<http://parks.seoul.go.kr/worldcuppark>

대표키워드

평화의공원, 하늘공원, 노출공원, 노출캠핑장, 난지천공원

정보수정 제안

좋아요 22

우리동네 바퀴 가게 알려주면, Pay 포인트가 팡팡!

최근 수정 2017.04.28

상세정보

포스트 리뷰 6

블로그·카페 리뷰 11,230

총 11,230건

월드컵공원 서울여니지드림센터-유아휴게실(수유실), 북카페

블로그

2017.04.11

안녕하세요 아델라입니다~ 지난 월요일에 저희 가족은 공원 나들이 다녀왔는데요~ 아무 리도 아기가 찻다보니 또 수유실, 유아휴게실 검색을ㅋㅋㅋ 갑자기 증가하면 기저귀 갈아 줘야 하고 이유식도 더워 먹어야하고 아직 모유수유 중이므로 맘마도 먹여야 하고요~ 그 러서 아기 데리고는 거의 디행...

<http://blog.naver.com/kyluvjh/220980119858> | ♥박두림, 김배죽, 김보라♡

2017 38회 롯데어린이미술대회 접수했어요

블로그

2017.04.05

어제 롯데백화점 근처에 놀일이 있어서 간 김에 88회 롯데 어린이 미술대회 접수하고 왔습 니다 어떻게 해야 하는지 궁금하신 이웃님들에게 정보를 드리고자 글을 작성합니다 안타 깁게도 온라인 접수가 안되니 직접 가까운 점포를 방문하여 사진 접수를 해야 하는데요 그 경도 선착순으로...

<http://blog.naver.com/youshin91/220975927801> | 뽀빠의 재밌는 디자인 이야기

서울분꽃나들이 명소 상암월드컵공원 [서울대이트코스추천명소]

블로그

2017.04.02

서울분꽃나들이 명소 상암월드컵공원 [서울대이트코스추천명소] 서울시 상암동 월드컵공 원은 상암월드컵 경기장 주변에 2002년 월드컵을 대비하여 조성한 공원이다 총 57의 테 마를 띠고 가지고 있는 공원으로 평화의공원과 가을 이면 역사축제가 열리는 하늘공원과 캠핑 장이있는 노출공원과...

그림 4-2. 네이버 지도의 ‘월드컵공원’ 검색 결과 상세보기 화면(PC 버전)

- 54 -



그림 4-3. 네이버 지도 ‘서울 맛집’
검색 결과(모바일 버전)

다음 지도의 경우, 검색 결과에 따른 POI 목록, 사용자가 남긴 평점, 장소평 그리고 관련 리뷰(티스토리, 네이버 블로그 포스트 등)를 제공한다(그림 4-4, 그림 4-5, 그림 4-6). 그러나 사용자가 직접 남긴 평점과 장소평의 개수가 많지 않고, 관련 리뷰에 직접 접속하여 내용을 파악해야 하는 어려움이 있다. 따라서 본 연구에서 제안하는 장소 선호도 분석 기법을 통해 수많은 관련 리뷰들의 텍스트를 분석하여 선호도를 도출한다면 부족한 평점 및 장소평 데이터를 보완할 수 있을 것이다.

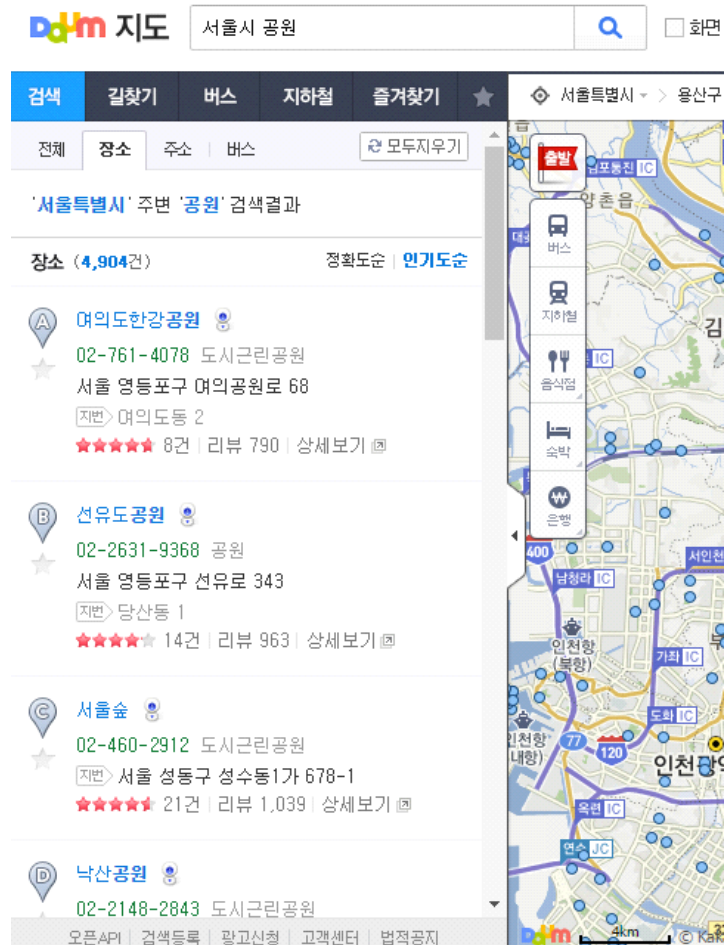
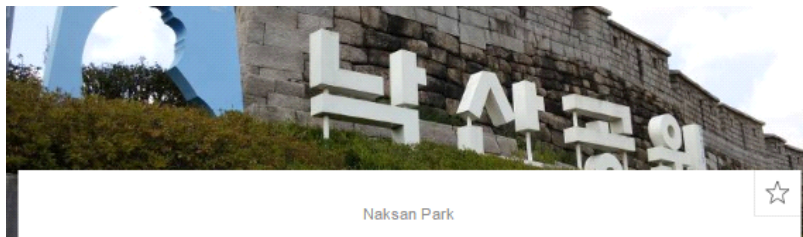


그림 4-4. 다음 지도의 검색결과 화면



Naksan Park



낙산공원

도시근린공원 · 평점 4.3 (3) · 리뷰 680



지도



길찾기



로드뷰



공유하기

평가하기 3

4.3 점 ★★★★★

★★★★★ 0/5 평가해주세요.

평가와 관련한 댓글, 악플은 삭제될 수 있습니다.

0 / 200

등록

★★★★★ 4

많이 알려져서 가는 길목에 찾집 풀비, 한가할때가 더 좋았음. 그래도 얼마안되는 서울의 중심가의 녹지공원, 꼭대기에서 일몰보기 좋음, 모기있음.

👤 좋아요 0 수가 | 2016.10.04. | 신고

★★★★★ 4

즐러가기좋아요

👤 좋아요 0 안녕— | 2013.05.18. | 신고

★★★★★ 5

서울의 숨은 데이트 장소 단 오르막 싫어하는 사람에게는 다소 비추. 올라간 것에 비해 경치가 많이 좋아요~ 소음도 없고요.

👤 좋아요 0 유광 | 2011.12.08. | 신고

그림 4-5. 다음 지도 ‘낙산공원’ 검색 결과의 평점 및 장소평

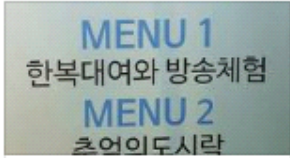








 <p>MENU 1 한복대여와 방송체험 MENU 2 추억의도시락</p> <p>2TV 생생정보, 경복궁 야간 특별관람 입장료..엿... 2TV 생생정보, 경복궁 야간 특별관람 입장료..엿날 도시락 소셜증 마먹방스튜디오 김영희 개봉... 누구인가 2017.05.02.</p>	 <p>대학로 낙산공원 벚꽃 ♥ 봄나들이 데이트코스</p> <p>대학로 낙산공원 벚꽃 구경하며 등산 같은 산책하기! ㅎㅎ봄나들이 데이트 코스♥ 저희 커플이... Rose Story♥ 2017.04.26.</p>	 <p>한양도성 순성길걷기(2) 장수마을 - 낙산공원</p> <p>장수마을에서 낙산공원까지 가는 길은 완전 오르막길이었습니 다 영화의 다음 장면을 알지 못... 내 삶을 가꾸는... 2017.04.25.</p>
 <p>혜화 놀거리_벽화마을 낙산공원</p> <p>오늘도 어김없이 등장하는 병아리 친구와 함께혜화에 있는 이화동 벽화마을 하고낙산공원 다녀.. 트끼가 트끼트... 2017.04.21.</p>	 <p>서울 갈만한곳 서울성곽길(한양도성길)따라 낙...</p> <p>주말에 가족과 함께 서울 갈만한곳! 서울성곽길 걸기 즐겨워~ 서울성곽길(한양도성길)은 남산 ... 해피맘 북북! 등.. 2017.04.20.</p>	 <p>[데이트코스/야경좋은곳] 동대문 낙산공원다녀왔...</p> <p>잇님들 안녕하세요~^^써늘이 예요~~오늘은 제가 소개해드릴 곳은동대문 낙산 공원이예요^^... 써늘의 배급배... 2017.04.20.</p>
 <p>나의 첫 출사 - 낙산공원, 이화벽화마을</p>	 <p>낙산공원★A6000으로 바라보다!</p>	 <p>낙산공원 서울성곽 봄풍 경</p>

그림 4-6. 다음 지도 ‘낙산공원’ 검색 결과의 관련 리뷰 목록

5. 결론

SNS 사용이 증가함에 따라 사용자들이 SNS를 통해 작성한 일상의 경험과 정보는 다른 사용자들의 의사결정에 영향을 미칠 수 있다. 따라서 많은 양의 SNS 데이터를 분석하여 의미 있는 정보를 도출하는 연구가 이루어지고 있으며, 본 연구에서는 SNS 텍스트를 이용하여 감성분석을 수행하고 장소 선호도 산출식을 제안하여 비정형의 텍스트 데이터로부터 장소에 대한 선호도를 도출하는 방법을 제안하였다. 이는 구글맵, 다음 지도에서 제공하는 기존의 장소에 대한 별점 정보와 달리 SNS 텍스트, 블로그 포스트, 뉴스기사 댓글과 같은 비정형의 텍스트 데이터로부터 장소 선호도를 도출할 수 있다는 점에서 의의가 있다.

이에 본 연구는 SNS 텍스트를 이용하여 감성분석을 수행하고 장소 선호도를 분석하기 위하여 장소를 표현하는 데에 적합한 감성어로 이루어진 공간 감성어 사전을 활용하고, SNS 텍스트의 특성을 고려하기 위하여 한글 초성 및 이모티콘을 감성표현으로써 감성분석의 대상으로 고려하였다. 또한, SNS 텍스트를 이용한 감성분석 결과를 바탕으로 장소 선호도를 계량화하는 장소 선호도 산출식을 제안하였다. 이는 비정형의 SNS 텍스트 정보로부터 사람들이 특정 장소에 대해 선호하는 정도를 수치로 나타내어 사용자들에게 유의미한 정보를 제공할 수 있다.

또한, 약 7개월 간의 트위터 데이터를 이용하여 서울 테마공원 36곳의 선호도를 분석하였다. 그 결과, 올림픽공원, 롯데월드, 도산공원, 동대문 역사문화공원, 여의도한강공원, 어린이대공원, 월드컵공원, 북서울꿈의숲, 남산공원 순으로 선호도가 높게 나타났다. 반면, 광나루자전거공원, 몽마르뜨공원, 서울창포원, 천호공원, 평화의공원, 한강랜드, 용마폭포공원은 트위터 텍스트 내에서의 언급빈도수가 적고, 긍정의 감성 또한 존재하지 않거나 적어 선호도가 낮게 나타났다. 또한, 트위터 텍스트에 대한 감성분석 수행 결과의 정확도를 평가하기 위하여 설문조사를 수행하였으며, 약 78.26%의 정확도(overall accuracy)를 보여 본 연구의 감성분석 결과를 바탕으로 한 장소 선호도 분석이 유의미함을 알 수 있었다.

본 연구에서 제안하는 SNS 텍스트를 이용한 감성분석 기반의 장소 선호도 분석기법은 트위터와 같은 SNS 텍스트뿐 아니라 일상에서 발생하는 블로그, 카페, 뉴스 기사 댓글 등의 비정형의 텍스트 데이터에 모두 적용이 가능하며, 현재 서비스 되고 있는 장소 정보 제공 서비스에 이를 활용하면 선호도를 기반으로 한 장소 정보를 제공할 수 있을 것이다.

본 연구의 한계는, SNS의 특성상 띄어쓰기를 하지 않거나 오타가 많이 발견되어 형태소 분석에 있어 정확성이 떨어지는 문제가 있었다. 이는 감성분석의 정확도를 떨어뜨리는 요인으로 추후 이에 대한 연구가 필요하다. 또한, 텍스트 내의 문맥을 고려하기 위하여 접속사, 부사 등의 보다 세밀한 언어학적 접근과 구문 분석에 대한 연구가 필요할 것으로 보인다. 뉴스기사 글과 달리 주어, 목적어, 서술어의 구조가 잘 갖추어지지 않은 SNS 텍스트의 특성을 고려하여 문맥을 파악하기 위한 심화 연구가 필요하다. 또한, 본 연구에서는 테마공원을 대상으로 하였지만, 분석하고자 하는 도메인에 따라 적합한 감성어 사전을 구축하고, 다양한 어휘 자원을 갖출 수 있도록 확장하는 연구가 필요하다. 마지막으로, 본 연구에서는 사전을 기반으로 감성분석을 수행하였으나, 추후에는 SVM 등과 같은 머신러닝 알고리즘 더 나아가서는 CNN 등의 딥러닝 알고리즘을 이용하여 대용량의 SNS 텍스트의 감성 분류를 자동화하고, 감성 분류 정확도를 향상시키는 연구가 필요하다. 이를 위해서는 양질의 많은 양의 SNS 텍스트를 수집하고, SNS 텍스트의 특성을 고려하여 학습을 위한 자질을 추출해야 한다. 즉, 부정확한 띄어쓰기와 오타 그리고 이모티콘과 신조어를 포함하고 있는 SNS 텍스트의 특성을 반영하여 음절 단위 나아가서는 음소 단위의 편집거리를 적용한 자질 추출을 통해 SNS 텍스트의 감성 분류 정확도를 향상시키는 연구가 선행되어야 한다.

참 고 문 헌

- 권세혁, 2007, 요인 분석과 신뢰도 계수, 한남대학교,
http://wolfpack.hnu.ac.kr/Book/SURVEY/fa_cron_surevy_wolfpack.pdf, pp.151-168.
- 김동영, 박제원, 최재현, 2014, SNS와 뉴스기사의 감성분석과 기계학습을 이용한 주가예측 모형 비교 연구, 한국 IT 서비스학회지, 제13권, 제3호, pp.211-233.
- 김용성, 전상, 전민아, 2014, 실시간 SNS 트렌드 분석에 기반한 여행지 추천 기법, 한국정보과학회 학술발표논문집, pp.291-293.
- 김윤화, 2015, SNS(소셜네트워크서비스) 이용추이 및 이용행태 분석, 정보통신정책연구원, 제16권, 제7호, pp.1-9.
- 남민지, 이은지, 신주현, 2015, 인스타그램 해시태그를 이용한 사용자 감정 분류 방법, 멀티미디어학회 논문지, 제18권, 제11호, pp.1391-1399.
- 박인조, 민경환, 2005, 한국어 감정단어의 목록 작성과 차원 탐색, 한국심리학회지: 사회 및 성격, 제19권, 제1호, pp.109-129.
- 박선우, 강금희, 고은정, 김은이, 2014, 멀티 모달 기반의 트위터 사용자 감성 흐름 분석, 한국정보과학회 학술발표논문집, pp.803-805.
- 명재석, 이동주, 이상구, 2008, 반자동으로 구축된 의미 사전을 이용한 한국어 상품평 분석 시스템, 정보과학회논문지: 소프트웨어 및 응용, 제35권, 제6호, pp.392-403.

- 박재희, 강영옥, 2014, 트윗을 이용한 서울시 주거환경 만족의 공간적 특성 분석-도시정책지표 보완을 위한 활용방안 모색, 한국도시지리학회지, 제17권, 제1호, pp.43-56.
- 사공원, 하성호, 박경배, 2016, 온라인 후기에 내재된 고객의 감성분석과 LQI 차원별 호텔서비스 품질 평가, 정보시스템연구, 제25권, 제3호, pp.217-245.
- 어정연, 2013, 장소가치의 형성요소가 장소선호도에 미치는 영향에 관한 연구, 국토지리학회지, 제47권, 제2호, pp.169-181.
- 이윤주, 서지훈, 최진탁, 2014, SNS 텍스트 콘텐츠를 활용한 오피니언마닝 기반의 패션 트렌드 마케팅 예측 분석, 한국정보기술학회논문지, 제12권, 제12호, pp.163-170.
- 임좌상, 김진만, 2014, 한국어 트위터의 감정 분류를 위한 기계학습의 실증적 비교, 멀티미디어학회논문지, 제17권, 제2호, pp.232-239.
- 양민혁, 정인선, 김용채, 2014, SNS 데이터를 활용한 국내대학 인식 및 선호도 분석, 한국빅데이터서비스학회 논문지, 제1권, 제1호, pp.1-13.
- 오효정, 윤보현, 최남현, 유철중, 2014, 소셜 빅데이터 내용 분석 기반 사용자 그룹별 선호지역 및 이동패턴 시각화, 한국정보기술학회논문지, 제12권, 제12호, pp.195-203.
- 안정국, 김희웅, 2015, 집단지성을 이용한 한글 감성어 사전 구축, 한국경영정보학회 학술대회논문집, pp.527-532.

- 임준엽, 황병연, 2015, 트위터 기반의 실시간 이벤트 지역 탐지 시스템, 정보처리학회논문지 소프트웨어 및 데이터 공학, 제4권, 제8호, pp.301-308.
- 오평화, 황병연, 2016, 트위터의 감정 분석을 통한 실시간 장소 추천 시스템, 한국전자거래학회지, 제21권, 제3호, pp.15-28.
- 이영민, 권필, 유기윤, 김지영, 2017, 한국어 장소 리뷰를 이용한 공간 감성어 사전 구축, 한국지형공간정보학회지, 제25권, 제2호, pp.3-12.
- 장재영, 2009, 온라인 쇼핑몰의 상품평 자동분류를 위한 감성분석 알고리즘, 한국전자거래학회지, 제14권, 제4호, pp.19-33.
- 장경애, 박상현, 김우제, 2015, 인터넷 감정기호를 이용한 긍정/부정 말뭉치구축 및 감정분류 자동화, 한국정보과학회, 제42권, 제4호, pp.512-521.
- 전소영, 신원용, 2015, 트위터에서 K-평균 군집화를 사용한 관심지역정보와 관광명소와의 관계, 한국통신학회 학술대회논문집, pp.270-271.
- 최진우, 배진수, 김효석, 2015, SNS 태그 분석 기반의 계절별 여행지 추천 기법, 한국통신학회 학술대회논문집, pp.499-500.
- 홍초희, 김학수, 2012, 트윗 감정 분류를 위한 다양한 기계학습 자질에 대한 비교 연구, 한국콘텐츠학회논문지, 제12권, 제12호, pp.471-478.

- 홍진주, 김세한, 박제원, 2016, 감성분석과 SVM 을 이용한 인터넷 악성 댓글 탐지 기법, 한국정보통신학회논문지, 제20권, 제2호, pp.260-267.
- Agarwal, A., Xie, B., Vovsha, I., Rambow, O., and Passonneau, R., 2011, Sentiment analysis of twitter data, In Proceedings of the workshop on languages in social media, Association for Computational Linguistics, pp.30-38.
- Denecke, K., 2008, Using sentiwordnet for multilingual sentiment analysis, Data Engineering Workshop, ICDEW 2008, IEEE 24th International Conference on IEEE, pp.507-512.
- Hridoy, S. A. A., Ekram, M. T., Islam, M. S., Ahmed, F., and Rahman, R. M., 2015, Localized twitter opinion mining using sentiment analysis, Decision Analytics, Vol.2, No.1, pp.1-19.
- Kouloumpis, E., Wilson, T., and Moore, J. D., 2011, Twitter sentiment analysis: The good the bad and the omg!, ICWSM, pp.538-541.
- Kolchyna, O., Souza, T. T., Treleaven, P., and Aste, T., 2015, Twitter sentiment analysis: lexicon method, machine learning method and their combination, arXiv preprint arXiv:1507.00955.
- Lu, Y., Castellanos, M., Dayal, U., and Zhai, C., 2011, Automatic construction of a context-aware sentiment lexicon: an optimization approach, In Proceedings of the 20th international conference on World wide web, ACM, pp.347-356.

- Nasukawa, T., and Yi, J., 2003, Sentiment analysis: Capturing favorability using natural language processing, In Proceedings of the 2nd international conference on Knowledge capture, ACM, pp.70-77.
- Smrž, P., 2006, Using WordNet for opinion mining, In Proceedings of the third international WordNet conference, pp.333-335.
- Tumasjan, A., Sprenger, T. O., Sandner, P. G., and Welpe, I. M., 2010, Predicting elections with twitter: What 140 characters reveal about political sentiment, ICWSM, Vol.10, No.1, pp.178-185.

부 록

- <부록 A> 공간 감성어 사전
- <부록 B> 추가 감성어 DB
- <부록 C> 서울 테마공원 POI를 포함하는
트위터 텍스트의 감성분석 결과
- <부록 D> 감성분석 결과에 대한 설문조사 응답 결과

<부록 A> 공간 감성어 사전

표 A-1. 공간 감성어 사전 목록(SSLex_single)

Table Name	SSLex_single						
ID	SP	Pt	Pb	ID	SP	Pt	Pb
S_1	추천/NC	1	1	S_58	만나/NC	0	0.612245
S_2	무료/NC	1	1	S_59	모이/NC	0	0.6
S_3	굿/NC	1	1	S_60	바라보/PV	0	0.673913
S_4	최고/NC	1	1	S_61	촬영/NC	0	0.653061
S_5	만족/NC	1	1	S_62	활동/NC	0	0.632653
S_6	감탄/NC	1	1	S_63	좋/PA	1	1
S_7	강추/NC	1	1	S_64	크/PA	0	0.542857
S_8	대박/NC	1	0.98	S_65	아름답/PA	1	1
S_9	실망/NC	-1	1	S_66	가깝/PA	1	1
S_10	유료/NC	0	0.764706	S_67	넓/PA	1	0.988235
S_11	인정/NC	1	0.969697	S_68	편하/PA	1	1
S_12	짱/NC	1	1	S_69	싸/PA	1	0.896104
S_13	천국/NC	1	0.978261	S_70	괜찮/PA	1	0.98374
S_14	최상/NC	1	1	S_71	맑/PA	1	1
S_15	최악/NC	-1	0.961538	S_72	비싸/PA	-1	0.967742
S_16	최적/NC	1	1	S_73	예쁘/PA	1	1
S_17	조용/NC	0	0.557692	S_74	작/PA	0	0.567376
S_18	새롭/PA	1	1	S_75	멋지/PA	1	1
S_19	소중/NC	1	1	S_76	새롭/PA	1	1
S_20	괜찮/PA	1	0.978723	S_77	놀람/PA	0	0.519084
S_21	괴롭/PA	-1	1	S_78	다르/PA	0	0.801587
S_22	괴상/NC	-1	0.976744	S_79	뛰어나/PA	1	1
S_23	그릴듯하/PA	1	0.84	S_80	멋있/PA	1	0.993007
S_24	높/PA	0	0.804878	S_81	색다르/PA	1	0.981132
S_25	답답/NC	-1	0.980392	S_82	습하/PA	-1	0.974359
S_26	따뜻/NC	1	1	S_83	아쉽/PA	-1	0.990476
S_27	딱딱하/PA	-1	1	S_84	엄청나/PA	1	0.990385
S_28	무난/NC	0	0.588235	S_85	즐겁/PA	1	1
S_29	번잡/NC	-1	0.978723	S_86	귀하/NC	1	0.974026
S_30	복잡/NC	-1	0.974359	S_87	나쁘/PA	-1	0.992126
S_31	불편/NC	-1	1	S_88	남/PA	-1	0.990385
S_32	사마/NC	-1	1	S_89	낮/PA	1	0.975904
S_33	수려/NC	1	1	S_90	낮익/PA	0	0.744828
S_34	신기/NC	1	1	S_91	멀/PA	-1	0.978022
S_35	심심/NC	-1	1	S_92	묘하/PA	0	0.830986
S_36	싸늘/NC	-1	1	S_93	무섭/PA	-1	0.991667
S_37	아기자기/NC	1	1	S_94	싫/PA	-1	1
S_38	아늑/NC	1	1	S_95	밝/PA	1	0.983362
S_39	아담/NC	1	0.965517	S_96	비좁/PA	-1	0.992701
S_40	안전/NC	1	1	S_97	배어나/PA	1	0.991597
S_41	적당/NC	1	0.966667	S_98	시끄럽/PA	-1	0.975806
S_42	친절/NC	1	1	S_99	시원시원하/PA	1	1
S_43	쾌적/NC	1	1	S_100	알차/PA	1	1
S_44	특별/NC	1	1	S_101	좁/PA	-1	0.981308
S_45	평범/NC	0	0.730769	S_102	활기차/PA	1	0.992248
S_46	포근/NC	1	1	S_103	힘들/PV	-1	1
S_47	느끼/PV	0	0.794118	S_104	재미있/PV	1	1
S_48	산책/NC	1	1	S_105	상쾌/NC	1	0.992857
S_49	늘/PA	1	1	S_106	깨끗/NC	1	1
S_50	견/PV	0	0.66	S_107	즐기/PV	1	1
S_51	나들이/NC	1	1	S_108	배우/NC	1	0.987179
S_52	운동/NC	1	1	S_109	쉬/NC	1	1
S_53	돌아다니/PV	0	0.64	S_110	노닐/PV	1	0.979167
S_54	취하/PV	1	0.964286	S_111	몸비/PA	0	0.47651
S_55	트리/PV	1	1	S_112	복제대/PV	0	0.47973
S_56	힐링하/NC	1	0.979592	S_113	손꼽/PV	1	1

Table Name	SSLex_single						
ID	SP	Pt	Pb	ID	SP	Pt	Pb
S_57	뛰/PV	0	0.765957	S_114	신나/NC	1	1

표 A-2. 공간 감성어 사전 목록(SSLex_combi)

Table Name	SSLex_combi									
ID	SS	SP	Pt	Pb	ID	SS	SP	Pt	Pb	
C_1	역/NC	가깝/PA	1	1	C_53	허브향/NC	취하/PV	1	0.964286	
C_2	공원/NC	편찮/PA	1	0.978723	C_54	휴식/NC	취하/PV	1	0.964286	
C_3	조용/NC	편찮/PA	1	0.978723	C_55	공원/NC	크/PA	0	0.542857	
C_4	장소/NC	넓/PA	1	0.988235	C_56	규모/NC	크/PA	0	0.542857	
C_5	주차장/NC	넓/PA	1	0.988235	C_57	시설/NC	편하/PA	1	1	
C_6	향/NC	느끼/PV	0	0.794118	C_58	오르막/NC	많/PA	-1	0.97561	
C_7	여유/NC	느끼/PV	0	0.794118	C_59	학생/NC	많/PA	0	0.75	
C_8	담/NC	멋지/PA	1	1	C_60	꽃/NC	많/PA	1	0.97561	
C_9	멍값/NC	비싸/PA	-1	0.967742	C_61	구경거리/NC	많/PA	1	1	
C_10	물가/NC	비싸/PA	-1	0.967742	C_62	주차장/NC	많/PA	1	1	
C_11	분위기/NC	색다르/PA	1	0.981132	C_63	행사/NC	많/PA	1	0.870968	
C_12	주차비용/NC	싸/PA	1	0.896104	C_64	쉽터/NC	많/PA	1	1	
C_13	밤/NC	아름답/PA	1	1	C_65	음식점/NC	많/PA	1	1	
C_14	서울에서/NC	아름답/PA	1	1	C_66	운동시설/NC	많/PA	1	1	
C_15	야경/NC	아름답/PA	1	1	C_67	문화재/NC	많/PA	1	1	
C_16	이용하기/NC	어렵/PA	-1	1	C_68	관광객/NC	많/PA	0	0.431373	
C_17	갈대/NC	예쁘/PA	1	1	C_69	편의시설/NC	많/PA	1	0.978723	
C_18	단풍/NC	예쁘/PA	1	1	C_70	산책로/NC	많/PA	1	1	
C_19	야경/NC	예쁘/PA	1	1	C_71	나무/NC	많/PA	1	0.989247	
C_20	공원/NC	작/PA	0	0.567376	C_72	차/NB	많/PA	-1	0.890244	
C_21	주차장/NC	좁/PA	-1	0.981308	C_73	건물/NC	많/PA	0	0.763514	
C_22	가족/NC	좋/PA	1	1	C_74	상인/NC	많/PA	0	0.727891	
C_23	경치/NC	좋/PA	1	1	C_75	사람/NB	많/PA	0	0.684564	
C_24	공원/NC	좋/PA	1	1	C_76	어린이/NC	많/PA	0	0.653061	
C_25	나들이/NC	좋/PA	1	1	C_77	가족/NC	많/PA	0	0.652482	
C_26	나들이하기/NC	좋/PA	1	1	C_78	아이들/NC	많/PA	0	0.651007	
C_27	냄새/NC	좋/PA	1	1	C_79	상점/NC	많/PA	0	0.646259	
C_28	농구하기/NC	좋/PA	1	1	C_80	연인/NC	많/PA	0	0.610738	
C_29	데이트하기/NC	좋/PA	1	1	C_81	언덕/NC	많/PA	0	0.530612	
C_30	드라이브하기/NC	좋/PA	1	1	C_82	오르막/NC	적/PA	1	0.97561	
C_31	벤치/NC	좋/PA	1	1	C_83	학생/NC	적/PA	0	0.75	
C_32	분위기/NC	좋/PA	1	1	C_84	꽃/NC	적/PA	-1	0.97561	
C_33	사전/NC	좋/PA	1	1	C_85	구경거리/NC	적/PA	-1	1	
C_34	산책/NC	좋/PA	1	1	C_86	주차장/NC	적/PA	-1	1	
C_35	산책하기/NC	좋/PA	1	1	C_87	행사/NC	적/PA	-1	0.870968	
C_36	시설/NC	좋/PA	1	1	C_88	쉽터/NC	적/PA	-1	1	
C_37	아이들/NC	좋/PA	1	1	C_89	음식점/NC	적/PA	-1	1	
C_38	야경/NC	좋/PA	1	1	C_90	운동시설/NC	적/PA	-1	1	
C_39	여유/NC	좋/PA	1	1	C_91	문화재/NC	적/PA	-1	1	
C_40	운동하기/NC	좋/PA	1	1	C_92	관광객/NC	적/PA	0	0.431373	
C_41	이용/NC	좋/PA	1	1	C_93	편의시설/NC	적/PA	-1	0.978723	
C_42	일광욕하기/NC	좋/PA	1	1	C_94	산책로/NC	적/PA	-1	1	
C_43	일몰때/NC	좋/PA	1	1	C_95	나무/NC	적/PA	-1	0.989247	
C_44	저녁/NC	좋/PA	1	1	C_96	차/NB	적/PA	1	0.890244	
C_45	전망/NC	좋/PA	1	1	C_97	건물/NC	적/PA	0	0.763514	
C_46	접근성/NC	좋/PA	1	1	C_98	상인/NC	적/PA	0	0.727891	
C_47	주차시설/NC	좋/PA	1	1	C_99	사람/NB	적/PA	0	0.684564	
C_48	편의시설/NC	좋/PA	1	1	C_100	어린이/NC	적/PA	0	0.653061	
C_49	휴식/NC	좋/PA	1	1	C_101	가족/NC	적/PA	0	0.652482	

Table Name	SSLex_combi								
ID	SS	SP	Pt	Pb	ID	SS	SP	Pt	Pb
C_50	힐링하기/NC	종/PA	1	1	C_102	아이들/NC	적/PA	0	0.651007
C_51	여유/NC	즐거/PV	1	1	C_103	상점/NC	적/PA	0	0.646259
C_52	운동/NC	즐거/PV	1	1	C_104	연인/NC	적/PA	0	0.610738

<부록 B> 추가 감성어 DB

표 B-1. 추가 감성어 DB 목록

Table Name	추가 긍정 감성어	Table Name	추가 부정 감성어
ID	SW	ID	SW
P_1	많/PA	N_1	없/PA
P_2	공원/NC	N_2	돈/NC
P_3	아이들/NC	N_3	관리/NC
P_4	갈/PA	N_4	필요/NC
P_5	산책하기/NC	N_5	가/PV
P_6	타/PV	N_6	그려/PV
P_7	느끼/PV	N_7	안되/PA
P_8	시설/NC	N_8	해변/NC
P_9	들/PV	N_9	먹을걸/NC
P_10	생각/NC	N_10	물건들/NC
P_11	사람/NC	N_11	이런걸/NC
P_12	보/PV	N_12	전시/NC
P_13	임/PV	N_13	주제별/NC
P_14	크/PA	N_14	즐기/PV
P_15	놀/PV	N_15	칼빙수/NC
P_16	가족/NC	N_16	아프/PA
P_17	장소/NC	N_17	평가/NC
P_18	다양한/NC	N_18	거울/NC
P_19	좋/NC	N_19	기대/NC
P_20	맛있/PA	N_20	동물들이/NC
P_21	견/PV	N_21	블로그에서/NC
P_22	자전거/NC	N_22	맛/NC
P_23	꽃/NC	N_23	받/PV
P_24	보이/PV	N_24	서비스/NC
P_25	주변/NC	N_25	좋/PA
P_26	찾/PV	N_26	입장료/NC
P_27	근처/NC	N_27	정도/NC
P_28	만들/PV	N_28	곳/NC
P_29	매/NC	N_29	기다리/PV
P_30	작/PA	N_30	나오/PV
P_31	둘/PV	N_31	완전/NC
P_32	피/PV	N_32	진짜/NC
P_33	알/PV	N_33	되/PV
P_34	불/PV	N_34	손님/NC
P_35	여름/NC	N_35	추/NC
P_36	오르/PV	N_36	말/NC
P_37	가격/NC	N_37	망/NC
P_38	먹/PV	N_38	모르/PV
P_39	어린이/NC	N_39	음식점/NC
P_40	위하/PV	N_40	이렇/PA
P_41	음식/NC	N_41	그렇/PA
P_42	동물원/NC	N_42	맛없/PA
P_43	이용/NC	N_43	여자/NC
P_44	자연/NC	N_44	카운터/NC
P_45	공간/NC	N_45	가지/PV
P_46	아이/NC	N_46	국물/NC
P_47	집/NC	N_47	다리/NC
P_48	들리/PV	N_48	터덜/PA
P_49	살/PV	N_49	돌문화/NC
P_50	구경하/NC	N_50	마삼/NC
P_51	박물관/NC	N_51	명칭/NC
P_52	공연/NC	N_52	불친절/NC
P_53	보내/PV	N_53	빠지/PV
P_54	체험/NC	N_54	수정/NC
P_55	공/NC	N_55	시키/PV

Table Name	추가 긍정 감성어	Table Name	추가 부정 감성어
ID	SW	ID	SW
P_56	애/NC	N_56	아줌마/NC
P_57	이쁘/PA	N_57	은은하게/NC
P_58	다니/PV	N_58	장사/NC
P_59	아니/PA	N_59	직원/NC
P_60	경치/NC	N_60	차/NC
P_61	꾸미/PV	N_61	장사/NC
P_62	안/NC	N_62	궁금/NC
P_63	주차/NC	N_63	동내/NC
P_64	주차장/NC	N_64	산책/NC
P_65	추억/NC	N_65	성인/NC
P_66	코스/NC	N_66	식당/NC
P_67	사진/NC	N_67	일/NC
P_68	저렴한/NC	N_68	장비/NC
P_69	가볍/PA	N_69	전화/NC
P_70	들르/PV	N_70	개판/NC
P_71	맛나/PA	N_71	계곡/NC
P_72	산책로/NC	N_72	관광객/NC
P_73	과거/NC	N_73	누나/NC
P_74	사/PV	N_74	뜨/PV
P_75	옛날/NC	N_75	바람/NC
P_76	잡되/PV	N_76	바쁘/PA
P_77	좋습니/NC	N_77	별/NC
P_78	미로/NC	N_78	불거리/NC
P_79	분수/NC	N_79	비/NC
P_80	마음/NC	N_80	소리/NC
P_81	바다/NC	N_81	아깝/PA
P_82	분위기/NC	N_82	어렵/PA
P_83	쉬/PV	N_83	영양/NC
P_84	야경/NC	N_84	예전/NC
P_85	입/PV	N_85	지저분함/NC
P_86	풍경/NC	N_86	참고/NC
P_87	행사/NC	N_87	커피/NC
P_88	규모/NC	N_88	하/PV
P_89	날/NC	N_89	한마디/NC
P_90	제미/NC	N_90	한번정도는/NC
P_91	가능/NC	N_91	호텔/NC
P_92	테리/PV	N_92	객실/NC
P_93	어른들/NC	N_93	거리/NC
P_94	여유/NC	N_94	곰팡이/NC
P_95	기차/NC	N_95	구하/PV
P_96	따르/PV	N_96	그려러니/NC
P_97	운영/NC	N_97	그저그렇다/NC
P_98	있/NC	N_98	기구/NC
P_99	하/PV	N_99	기본/NC
P_100	가기좋은/NC	N_100	기분/NC
P_101	되새기/PV	N_101	길/NC
P_102	모습/NC	N_102	냄새/NC
P_103	선생님/NC	N_103	널/PV
P_104	애기/NC	N_104	널/PV
P_105	저녁/NC	N_105	다르/PA
P_106	추억하게/NC	N_106	담/PV
P_107	포탈/NC	N_107	동물들/NC
P_108	계절/NC	N_108	많이오는데/NC
P_109	나무/NC	N_109	말씀드리/PV
P_110	바이킹/NC	N_110	문제/NC
P_111	알/NC	N_111	밖/NC
P_112	운동하기/NC	N_112	반찬/NC
P_113	술/NC	N_113	밖/NC
P_114	열/NC	N_114	별도많고/NC
P_115	입장/NC	N_115	보수공사/NC

Table Name	추가 긍정 감성어	Table Name	추가 부정 감성어
ID	SW	ID	SW
P_116	종류/NC	N_116	불개없고/NC
P_117	놀이기구/NC	N_117	부족/NC
P_118	지나/PV	N_118	불친/NC
P_119	휴식/NC	N_119	불쾌/NC
P_120	힘들/PV	N_120	비하/PV
P_121	개방/NC	N_121	사장님/NC
P_122	구경/NC	N_122	서방/NC
P_123	깨끗한/NC	N_123	설명/NC
P_124	도시/NC	N_124	성의/NC
P_125	돛자리/NC	N_125	심하/PA
P_126	산/NC	N_126	얕/PV
P_127	아침/NC	N_127	어/NC
P_128	장관/NC	N_128	업소/NC
P_129	좋은곳/NC	N_129	여기/PV
P_130	차/NC	N_130	오분자기/NC
P_131	호수/NC	N_131	외국/NC
P_132	관람/NC	N_132	원장님/NC
P_133	기구/NC	N_133	유치원/NC
P_134	깜깜/NC	N_134	자/PV
P_135	데이트/NC	N_135	자리/NC
P_136	수영장/NC	N_136	잘못/NC
P_137	위치한/NC	N_137	잠/NC
P_138	지역/NC	N_138	잠/PV
P_139	찍/PV	N_139	재미도/NC
P_140	가족들/NC	N_140	절대가지마/NC
P_141	눈/PA	N_141	정신/NC
P_142	멀/PA	N_142	주문/NC
P_143	벚꽃/NC	N_143	지나치/PV
P_144	상대적/NC	N_144	지배인/NC
P_145	속/NC	N_145	직원들/NC
P_146	습지/NC	N_146	짜증/NC
P_147	시원/NC	N_147	쭈크/NC
P_148	식물원/NC	N_148	처음/NC
P_149	의미/NC	N_149	친구/NC
P_150	정류장/NC	N_150	테이블/NC
P_151	醬/PA	N_151	트/PV
P_152	코끼리/NC	N_152	팔/PV
P_153	가을/NC	N_153	평일/NC
P_154	날씨/NC	N_154	할곳이/NC
P_155	놀이/NC	N_155	형편/NC
P_156	돌테길/NC	N_156	화장실/NC
P_157	무섭/PA	N_157	가방/NC
P_158	바깥/NC	N_158	가성비/NC
P_159	방문/NC	N_159	감사/NC
P_160	역사/NC	N_160	값/NC
P_161	조성/NC	N_161	강요/NC
P_162	걸/PV	N_162	강요함/NC
P_163	교육/NC	N_163	개뿔/NC
P_164	굿모닝/NC	N_164	개통/NC
P_165	근교/NC	N_165	거금/NC
P_166	꼭대기/NC	N_166	거지/NC
P_167	내리/PV	N_167	걸리/PV
P_168	네요/NC	N_168	계산/NC
P_169	동문/NC	N_169	계산대/NC
P_170	둘러보/PV	N_170	고기/NC
P_171	많/NC	N_171	고르/PV
P_172	물/NC	N_172	고용/NC
P_173	서틀버스/NC	N_173	곳곳/NC
P_174	심/NC	N_174	공짜/NC
P_175	오후/NC	N_175	관람객/NC

Table Name	추가 긍정 감성어	Table Name	추가 부정 감성어
ID	SW	ID	SW
P_176	운행/NC	N_176	괜찮/PA
P_177	음악/NC	N_177	국민/NC
P_178	전망/NC	N_178	굴/PV
P_179	전망대/NC	N_179	그날/NC
P_180	주차하면/NC	N_180	기르/PV
P_181	카페/NC	N_181	깨끗/NC
P_182	캠핑장/NC	N_182	끓이/PV
P_183	파악해/NC	N_183	나가/PV
P_184	각종/NC	N_184	나열/NC
P_185	강아지/NC	N_185	날/NC
P_186	깊/PA	N_186	낮/NC
P_187	나무들/NC	N_187	넌킨/NC
P_188	넌/PV	N_188	너어무/NC
P_189	돌아다니/PV	N_189	넓/PA
P_190	동물/NC	N_190	노년층/NC
P_191	밤/NC	N_191	누르/PV
P_192	사랑/NC	N_192	눈/NC
P_193	숲속/NC	N_193	단점/NC
P_194	시간/NC	N_194	단품/NC
P_195	아파트/NC	N_195	달/PA
P_196	연꽃/NC	N_196	대답/NC
P_197	열리/PV	N_197	덜컹거리/NC
P_198	절/NC	N_198	도로/NC
P_199	정문/NC	N_199	돈독/NC
P_200	조용/NC	N_200	둘리/PV
P_201	직업/NC	N_201	동굴카페/NC
P_202	편리/NC	N_202	들었/NC
P_203	건물/NC	N_203	똥물/NC
P_204	단풍/NC	N_204	마인드/NC
P_205	문/NC	N_205	말쟁하더니/NC
P_206	시원한/NC	N_206	말하라하니/NC
P_207	위치/NC	N_207	맛이없/NC
P_208	음악분수/NC	N_208	매니저님/NC
P_209	이상/NC	N_209	모/NC
P_210	입구/NC	N_210	목소리/NC
P_211	지하/NC	N_211	못먹을/NC
P_212	축제/NC	N_212	못미치/PA
P_213	치/PV	N_213	못자고/NC
P_214	테마파크/NC	N_214	무시/NC
P_215	할인/NC	N_215	물건/NC
P_216	놀이공원/NC	N_216	물놀이/NC
P_217	놀이터/NC	N_217	물리치료/NC
P_218	용/NC	N_218	뭔가/NC
P_219	유적/NC	N_219	미숫가루/NC
P_220	유지/NC	N_220	바닥/NC
P_221	이후/NC	N_221	배개/NC
P_222	장미/NC	N_222	병원/NC
P_223	처럼/NC	N_223	보트/NC
P_224	한눈/NC	N_224	본점/NC
P_225	호선/NC	N_225	부르/PV
P_226	간단한/NC	N_226	부탁드리/PV
P_227	감동/NC	N_227	분재원/NC
P_228	그늘막/NC	N_228	불/NC
P_229	떨어지/PV	N_229	불구/NC
P_230	미술관/NC	N_230	불친절한/NC
P_231	시작/NC	N_231	불편하게/NC
P_232	신나/PV	N_232	불편함/NC
P_233	어리/PA	N_233	비슷한/NC
P_234	예약/NC	N_234	빌리/PV
P_235	이해/NC	N_235	사과/NC

Table Name	추가 긍정 감성어	Table Name	추가 부정 감성어
ID	SW	ID	SW
P_236	인터넷/NC	N_236	사라지/PV
P_237	자리/NC	N_237	사람들/NC
P_238	최대/NC	N_238	사먹으며/NC
P_239	폭포/NC	N_239	상식/NC
P_240	하루/NC	N_240	생각나/PV
P_241	하루종일/NC	N_241	설계/NC
P_242	거대한/NC	N_242	성/NC
P_243	굴/NC	N_243	세금/NC
P_244	남/PV	N_244	세트/NC
P_245	동상/NC	N_245	소리지르/PV
P_246	만원/NC	N_246	수준/NC
P_247	분재/NC	N_247	수질/NC
P_248	사진찍/PV	N_248	쉽/PA
P_249	서/PV	N_249	스프/NC
P_250	쓰/PV	N_250	식품매장/NC
P_251	아기/NC	N_251	썩었/NC
P_252	어우러지/PV	N_252	쓰레기/NC
P_253	역사적/NC	N_253	쌈/PV
P_254	잉꼬/NC	N_254	아깁/NC
P_255	자랑/NC	N_255	안주/NC
P_256	조깅/NC	N_256	안하/PV
P_257	종구요/NC	N_257	얇/NC
P_258	직원분/NC	N_258	얇고../NC
P_259	추가요금/NC	N_259	얇/NC
P_260	한바퀴/NC	N_260	어머니/NC
P_261	결제/NC	N_261	어지럽/PA
P_262	경기장/NC	N_262	얼굴/NC
P_263	구성/NC	N_263	없구/NC
P_264	김/NC	N_264	엘레베이터/NC
P_265	다음/NC	N_265	엘리베이터/NC
P_266	도심속/NC	N_266	연락/NC
P_267	들락날락하/PV	N_267	열/PV
P_268	떨기/NC	N_268	옆자리/NC
P_269	마지막/NC	N_269	오래/PA
P_270	모/NC	N_270	올리/PV
P_271	바라보/PV	N_271	요즘/NC
P_272	생기/PV	N_272	요청/NC
P_273	석양/NC	N_273	우매우/NC
P_274	어린이들/NC	N_274	우의/NC
P_275	연인/NC	N_275	운동기구/NC
P_276	오전/NC	N_276	울/PV
P_277	위터파크/NC	N_277	원하/PV
P_278	있습니/NC	N_278	위/NC
P_279	자동차/NC	N_279	유크/NC
P_280	잔디밭/NC	N_280	유모차/NC
P_281	전시품/NC	N_281	으로/NC
P_282	조아/NC	N_282	의자/NC
P_283	지키/PV	N_283	이글/NC
P_284	테니스장/NC	N_284	이야기거리/NC
P_285	펼치/PV	N_285	이용객/NC
P_286	품/PV	N_286	인당/NC
P_287	환경/NC	N_287	인디언/NC
P_288	건너/PV	N_288	자기들이/NC
P_289	공연장/NC	N_289	잔소리/NC
P_290	광장/NC	N_290	잘나/PA
P_291	국내/NC	N_291	잘되어있/NC
P_292	그늘/NC	N_292	장난/NC
P_293	눈/NC	N_293	적/PA
P_294	늘/PV	N_294	적합/NC
P_295	테이트하기/NC	N_295	전체적/NC

Table Name	추가 긍정 감성어	Table Name	추가 부정 감성어
ID	SW	ID	SW
P_296	동네/NC	N_296	절/PV
P_297	맞추/PV	N_297	정말별/NC
P_298	명소/NC	N_298	조합/NC
P_299	발전/NC	N_299	종업원들이/NC
P_300	범/NC	N_300	주방장/NC
P_301	불/NC	N_301	주인/NC
P_302	불편/NC	N_302	준비/NC
P_303	새/NC	N_303	지저분/NC
P_304	시원하게/NC	N_304	진입하기/NC
P_305	아래/NC	N_305	짜/PV
P_306	안타깝/PA	N_306	짜증나/PV
P_307	여행/NC	N_307	찜질방/NC
P_308	이루/PV	N_308	참/PV
P_309	인공/NC	N_309	창렬임/NC
P_310	진경/NC	N_310	최/NC
P_311	조망/NC	N_311	축소되서/NC
P_312	좁/NC	N_312	치중/NC
P_313	최고입니/NC	N_313	친절/NC
P_314	최근/NC	N_314	캠핑책/NC
P_315	커다랗/PA	N_315	타워/NC
P_316	평화/NC	N_316	타프값도/NC
P_317	학교/NC	N_317	태도/NC
P_318	함께하/PV	N_318	테이크아웃/NC
P_319	게임/NC	N_319	토끼/NC
P_320	경관/NC	N_320	통제/NC
P_321	고맙/PA	N_321	튜브/NC
P_322	고양이/NC	N_322	편의/NC
P_323	골목/NC	N_323	평상/NC
P_324	관광/NC	N_324	표정/NC
P_325	교통/NC	N_325	플라스틱컵/NC
P_326	길속/NC	N_326	하시길/NC
P_327	깎/PV	N_327	학원/NC
P_328	단위/NC	N_328	향의/NC
P_329	대하/PV	N_329	해물/NC
P_330	뒤편/PA	N_330	핸드폰/NC
P_331	떠나/PV	N_331	후/NC
P_332	만발/NC	N_332	미안/NC
P_333	맞/PV	N_333	그동안/NC
P_334	매력/NC	N_334	운동시설/NC
P_335	명당/NC	N_335	잠들/PV
P_336	모이/PV	N_336	폐업분위기/NC
P_337	바라/PV		
P_338	방문하기/NC		
P_339	버스/NC		
P_340	비/NC		
P_341	세계/NC		
P_342	술길/NC		
P_343	섬티/NC		
P_344	슬/PV		
P_345	얼/PV		
P_346	연결/NC		
P_347	연못/NC		
P_348	연일/NC		
P_349	위치해/NC		
P_350	이름/NC		
P_351	이유/NC		
P_352	접근성/NC		
P_353	접하/PV		
P_354	정원/NC		
P_355	조용한/NC		

Table Name	추가 긍정 감성어	Table Name	추가 부정 감성어
ID	SW	ID	SW
P_356	조형물/NC		
P_357	좋고./NC		
P_358	주민들/NC		
P_359	지하철/NC		
P_360	친절하시/NC		
P_361	필요한/NC		
P_362	한가운데/NC		
P_363	훌륭/NC		
P_364	힐링/NC		
P_365	먹거리/NC		
P_366	아쉽/PA		
P_367	정말/NC		
P_368	주말/NC		
P_369	봄/NC		

<부록 C> 서울 테마공원 POI를 포함하는 트위터 텍스트의 감성분석 결과

표 C-1. 서울 테마공원 POI를 포함하는 트위터 텍스트의 감성분석 결과

번호	POI	불용어 제거된 트위터 텍스트	공간 감성어 사전		이모티콘 고려		PT	NT
			주어	감성어	공간 감성어 및 추가 감성어	이모 티콘		
1	낙성대공원	도서관도 있고 좋네 (낙성대공원 () 서울특별시.)	낙성대공원	좋/PA	NA	NA	1	0
2	낙성대공원	시작 ... (낙성대공원 () 서울특별시.)	NA	NA	시작/NC	...	1	0
3	남산공원	내일은 우리 뽕썰근 데리고 산책해야지~ . 산책 공원 매미소리 좋다 강아지 개 남산공원 남산공원	NA	NA	산책/NC	~	1	0
4	남산공원	남산공원 처음으로 갔다. 길 몰라서 재미있었다 ㅋㅋ...	NA	NA	재미/NC	ㅋㅋ	1	0
5	노을공원	왜 나는 이런 날씨에만 녹화를 나오는가... 태안에서 눈보라 맞은 이후 정확히 년후 상암동 겨울왕국을 만나다. 노을공원	NA	NA	나오/PV	...	0	-1
6	노을공원	흐린 날씨... 돈키우스 트레킹 노을공원 노을공원	NA	NA	날씨/NC	...	0	-1
7	노을공원	아이 좋아~ 노을공원 캠핑장	노을공원	좋/PA	좋/PA	~	2	0
8	노을공원	급 번개 나들이:) (노을공원 서울특별시.)	노을공원	나들이/NC	나들이/NC	:)	2	0
9	도산공원	산책 (도산공원 () 서울특별시.)	도산공원	산책/NC	NA	NA	1	0
10	도산공원	멜팅샵 현대카드 고메워크 ... 바질 사랑하여 ㅋㅋㅋㅋ트트트 소프트웨어랩 카페레제파스타 장백기 왔다 ㅋㅋㅋㅋ트 도산공원...	NA	NA	보/PV	ㅋㅋ ㅋ	1	0
11	도산공원	오늘은 회사언니랑 도산공원까지?? 단풍구경~!!?ㅎㅎ 운동 겸 산책 압구정 도산공원 점심시간 쉬는시간 행복 단풍 가을 도산공원	도산공원	운동/NC	NA	NA	1	0
12	도산공원	도산공원 산책하다 도산공원 서울 강남 도산공원	도산공원	산책/NC	NA	NA	1	0
13	도산공원	헤헤헤 여긴 일부러 이렇게 레이어로 쌓아주는걸까 암튼 넘넘 맛있소 우린아리아께는 언제 가보는걸까? ㅋㅋㅋ 런치 오마카세 스시 스시선수 도산공원 도산공원맛집ㅋㅋㅋ	NA	NA	집/NC	ㅋㅋㅋ	1	0
14	도산공원	도산공원 카페 . 한적할때 더 좋은...	NA	NA	좋/PA	...	1	0
15	도산공원	도산공원 앞 커피지인 도피오 에스프레소터블 오랫동안 진한 에스프레소 맛나네...	NA	NA	맛나/PA	...	1	0
16	도산공원	바지락 매생이 비빔면 대박! 멜팅샵 도산공원	도산공원	대박/NC	NA	NA	0.98	0
17	도산공원	고 투더 헬!!!! 영혼의 안식처..... 모즈 신사동!! 모즈 신사동모즈 일식 갓포 첫번째 ?사이모즈 신사동 도산공원 해본 천국의 창고 ㅎㅎ ?사이 모즈	도산공원	천국/NC	NA	NA	0.9782 61	0

번호	POI	불용어 제거된 트위터 텍스트	공간 감성어 사전		이모티콘 고려		PT	NT
			주어	감성어	공간 감성어 및 추가 감성어	이모 티콘		
18	도산공원	생일쿠폰 으로 내가 참 좋아하는 티라미슈 로열자몽티 홀쭉. 역시 자몽티는 꽤 진하게 잘 만들어졌어. 조용하다 일요일 저녁 도산공원 아티제 아티제 도산공원점	도산공원	조용/NC	NA	NA	0	0
19	동대문역사 문화공원	아주 예쁘다~^^ 꽃전구 동대문역사문화공원	동대문역사문 화공원	예쁘/PA	예쁘/PA	~^^	2	0
20	동대문역사 문화공원	동대문 롯데피트인 컬처스타에 레이지수잔의 공간이 생겼어요!!:) 완벽하진 않지만 이번주내로 아급야급 채워나갈게요. 동대문역사문화공원역 - 번출구 사이에 바로 들어올수있는...	NA	NA	생기/PV	:)	1	0
21	동대문역사 문화공원	좀 이쁘긴 하더라~ 장미 꽃밭 동대문역사문화공원 동대문 디자인 플라자 장미 꽃밭	NA	NA	하/PV	~	1	0
22	롯데월드	환상에 나라 롯데월드???? 힘들다 ? 롯데월드 잠실 놀이기구 환상에나라 꿈과희망 어지러움 롯데월드	롯데월드	힘들/PV	NA	NA	0	-1
23	롯데월드	퇴근후 날아와서 잠실 이성당 이성당 잠실 호선 힘들어 롯데월드 ()	롯데월드	힘들/PV	NA	NA	0	-1
24	롯데월드	캘리포니아피자키친 피자 올거준맛크 빵이마이?! 배장이겁나한가해서그런가(..)서버분 이엄청리필여부물어파주셨다(....) 캘리포니아 피자키친 롯데월드몰점	NA	NA	맛/NC	ㅋ	0	-1
25	롯데월드	둘이서만 와보는 롯데월드 이케 얼마만인가 ㅋ 아틀란티스 점점인게 아쉽군... 환상과모험의나라 롯데월드	NA	NA	아쉽/PA	...	0	-1
26	롯데월드	롯데월드 한동안 못가? 느~~~무 힘들어 ㅋㅋㅋ (마지막쪽에 롯데월드 댄서들과 배경음악이 딱맞아ㅋㅋㅋ) 잠실 롯데월드 꿈과희망의나라 페스티벌...	NA	NA	힘들/PV	ㅋㅋㅋ	0	-1
27	롯데월드	월드시네마에서 서부전선 감상.. 중간중간은 한참 웃었는데... 내용은 다 예상되지만 생각없이 웃을수 있었네.. 항상 그렇지만 끝은 아쉬웠네... 잠실 롯데월드시네마	롯데월드	아쉽/PA	NA	NA	0	-0.99 048
28	롯데월드	실물 처음 봄 ㅋㅋㅋㅋㅋㅋㅋㅋ 신기하게 생겼다 ㅋㅋㅋ 롯데월드몰	롯데월드	신기/NC	봄/NC	ㅋㅋ ㅋㅋ ㅋㅋ	2	0
29	롯데월드	잠실의 메카 롯데월드 ... 멋지기는 하네... 안개가 아주짱 롯데월드 ()	롯데월드	멋지/PA	하/PV	...	2	0
30	롯데월드	운이 좋았어!! 제 롯데월드몰	롯데월드	좋/PA	NA	NA	1	0
31	롯데월드	요다 ㅋ 롯데월드몰 롯데시네마에서 처음으로 영화봄 ㅋ 다른건 다 좋은데 사운드가 병맛 (심지어 사운드 특화관인 슈퍼사운드? 관이었는데...) 입체감 전혀 없이 앞에서만 나와서...	NA	NA	봄/NC	ㅋ	1	0

번호	POI	불용어 제거된 트위터 텍스트	공간 감성어 사전		이모티콘 고려		PT	NT
			주어	감성어	공간 감성어 및 추가 감성어	이모 티콘		
32	롯데월드	지금 집에서 아침먹으며 보는 하늘이 꽤 아름답네 롯데월드 ()	롯데월드	아름답/PA	NA	NA	1	0
33	롯데월드	에스타그램 셀스타그램 팔스타그램 육스타그램 소통 육스타그램 육아 베이비그램 베베 베이비인스타 개월 롯데월드물 맞팔 좋아요 선팔 베이비 줌스타그램 육아...	롯데월드	좋/PA	NA	NA	1	0
34	롯데월드	: 금요일 서울 잠실 롯데월드 모험 과 신비가 가득한 나라 우리가 꿈꾸던 그곳 여기는 ~ 롯데월드 ~♡ 바이킹이랑 자이로드롭은...	NA	NA	늘/PV+L/E T	~~♡	1	0
35	롯데월드	마리메코 양말 유니코 유니코 양말은 언제나 기분이 좋아?????? 롯데월드물	롯데월드	좋/PA	NA	NA	1	0
36	롯데월드	롯데월드 세계 최대의 실내 놀이공원 롯데 롯데월드 놀이공원 서울 한국 대한민국 아이스링크 웃겨 좋아 사진 사진스타그램 테일리 맞팔 ...	사진/NC	좋/PA	NA	NA	1	0
37	롯데월드	? 힐링나들이 서울 롯데월드 꿈과희망보다는 비명과광란 사진은많은데 페북에올려버림 다들고마워용 ? 롯데월드 ()	롯데월드	나들이/NC	NA	NA	1	0
38	롯데월드	자본주의 최고야 (롯데월드물 ())	롯데월드	최고/NC	NA	NA	1	0
39	롯데월드	롯데월드 가든스테이지 깨끗하게! 맞게! 자신있게! ? 미소천사 ? 남시오 ~ ...	롯데월드	깨끗/NC	NA	NA	1	0
40	롯데월드	내일 시험이 있는데..... 롯데월드 크리스마스 분위기 너무 좋다 ?? 롯데월드 ()	분위기/NC	좋/PA	NA	NA	1	0
41	롯데월드	지난 월 마지막에 먹은 롯데월드 입장전에 먹은 아이스크림^^ 맛났다♥? 멕스타그램 멕스타그램 ...	NA	NA	맛나/PA	♥	1	0
42	롯데월드	난 여기 야채빵 단팥빵보다 크림 든 빵들이 맛있더라~ 이성당 롯데월드 ()	NA	NA	맛있/PA	~	1	0
43	롯데월드	재미난놀이 부메랑 움짤 운동화 롯데월드점 참 재미난게 많네 ㅎㅎ	NA	NA	많/PA	ㅎㅎ	1	0
44	롯데월드	쇼케이스 ?? 생각보다 완전 가깝다TT TT 얼른 보고싶어 사랑해 ?? 쇼케이스 롯데월드 롯데월드 ()	NA	NA	가깝/PA	TT TT	1	0
45	롯데월드	롯데월드 가든스테이지 그댄 먼 곳만 보네요~ 내가 바로 여기 있는데... (러브홀릭 '인형의 꿈' 중에서) ...	NA	NA	보/PV	~	1	0
46	롯데월드	지난 방학때 서울 롯데월드에서 가고 친구들과 함께 즐겁게 놀았다~	NA	NA	놀/PV	~	1	0
47	롯데월드	: 금요일 서울 잠실 롯데월드 완전 멋있었던 야간 퍼레이드... 놀이동산 로터리 럭스타그램 신바트의모험 후랜치레볼루션...	롯데월드	멋있/PA	NA	NA	0.9930 07	0
48	롯데월드	밝은너석 롯데월드 ()	롯데월드	밝/PA	NA	NA	0.9893 62	0
49	롯데월드	으와아아 넓어 롯데월드물	롯데월드	넓/PA	NA	NA	0.9882 35	0

번호	POI	불용어 제거된 트위터 텍스트	공간 감성어 사전		이모티콘 고려		PT	NT
			주어	감성어	공간 감성어 및 추가 감성어	이모 티콘		
50	롯데월드	드디어! 근데 생각보다 쾡쾡음ㄱ (롯데월드볼 ())	롯데월드	쾡쾡/PA	NA	NA	0.9812 32	0
51	롯데월드	월 일 종료. 오늘 하루 종일 두통 때문에 엄청 힘들었는데 마지막은 큰 선물을 받는 느낌. 롯데월드볼 길리안 초콜릿 카페 초콜릿빙수 이런 빙수 처음이야...! 달아...	롯데월드	느끼/PV	NA	NA	0	0
52	롯데월드	제 롯데월드 높다.? 일상 빌딩 팔로우 선풍 맞팔 맞팔해요 선풍해요 인천 잠실 송파구 서울 석촌호수 좋아요 제 롯데월드볼	롯데월드	높/PA	NA	NA	0	0
53	롯데월드	오늘의 점심은 높은데서. 도림 롯데월드호텔-도림	롯데월드	높/PA	NA	NA	0	0
54	롯데월드	이곳에서도 보일 정도로 제 롯데월드 가 높아졌다. 어릴 때에는 뉴욕 엠파이어와 빌딩이 최고인 줄만 알았는데 급성장한 중국에는 저런 게 수백 동이나 된다고 하고... ...	롯데월드	높/PA	NA	NA	0	0
55	롯데월드	열심히 걷고 있었음 (롯데월드 아쿠아리움 () 서울)	롯데월드	건/PV	NA	NA	0	0
56	몽마르트공원	비오기전 단풍 구경 몽마르트공원 상쾌한공기 토끼천국 서래마을 몽마르트 공원	몽마르트공원	천국/NC	NA	NA	0.9782 61	0
57	북서울꿈의숲	집사람이랑 밖에서 저녁 먹고 산책목~ㄱ 퍼몬치드 누려~~!!ㄱ (북서울꿈의숲 () -)	NA	NA	숲/NC	~~ㄱ ~~ㄱ	1	0
58	북서울꿈의숲	북서울꿈의숲 전망대 집에 오고서야 불꽃축제가 있었다는 걸 알게됨 알았으면 저 위에서 좀 더 기다려볼 걸 그래도 안보이려나 ㄱ 북서울 꿈의 숲	NA	NA	보/PV	ㄱ	1	0
59	북서울꿈의숲	예비둘찌형부(?)랑 둘쨈언니랑 엄마랑 함께... 두번째 뮤지컬문준경 보러 ㅎㅎ (북서울꿈의숲 아트센터 () 서울특별시.)	NA	NA	보/PV	ㅎㅎ	1	0
60	서서울호수공원	이런 !!! 팀 쿡이 잡스?의 유산 애플을 "삼성 따라쟁이"로 만들었다 사바... 쾡 하고 애플쾡 하고 다르게 뭐임? (서서울호수공원 서울특별시. 서울특별시.)	서서울호수공원	다르/PA	NA	NA	0	0
61	어린이대공원	사슴아 침은 묻히지 말고 먹으면 안되겠니... 어린이대공원 동물원 먹이주기 사슴 어린이대공원	NA	NA	안되/PA	...	0	-1
62	어린이대공원	간만에 날씨 좋아서 아들테리고 나들이. (어린이대공원 (') 서울특별시.)	어린이대공원, 나들이/NC	나들이/NC, 쾡/PA	NA	NA	2	0
63	어린이대공원	가족 나들이~ 어린이대공원에서 네식구 총출동~!!! (어린이대공원 동물원 (') 서울특별시.)	어린이대공원	나들이/NC	나들이/NC	~	2	0
64	여의도한강공원	얼마나 고마운지 모르겠어! 아무리 사진이 별로 좋지 않아도 가장 좋은 추억을 남겼다~ 히힛 카우 제인 하나 추억이라부르게 여의도한강공원	사진/NC	쾡/PA	NA	NA	0	-1

번호	POI	불용어 제거된 트위터 텍스트	공간 감성어 사전		이모티콘 고려		PT	NT
			주어	감성어	공간 감성어 및 추가 감성어	이모 티콘		
65	여의도한강공원	예고예고 땀 줄줄 힘들다. 여기서 여의도 가로지르면 집인데 그래도 안양천까지 가서 도립천타고 들어갈거다. 내일은 지각 확정이겠군 (여의도한강공원 ())	여기/MA	힘들/PV	NA	NA	0	-1
66	여의도한강공원	좋다~ (여의도한강공원 () - 서울특별시. 서울시)	여의도한강공 원	좋/PA	좋/PA	~	2	0
67	여의도한강공원	오늘 서울 하늘 맑음 여직원회식 여의도 한강 나들이 여의도한강공원	여의도한강공 원	나들이/NC	NA	NA	1	0
68	여의도한강공원	빼에로 아줌마. 몇년 전부터 주로 서울광장에서 보이시는 아주머니 간혹 서울숲 여의도한강공원 남산 등지에서도 보이던... (복장 + 얼굴칠 은 항상 동일 빨간코는...	NA	NA	늘/PV+L/E T	...	1	0
69	여의도한강공원	서강대교 아래 트랙 한바퀴 원리로 돌기 성공~~! 라이딩 자전거 파이어아이 여의도한강공원 서강대교 아래	NA	NA	공/NC	~~	1	0
70	여의도한강공원	이쁘다 ~ ~ ~ ~ ~ 바람살랑살랑부는 여의도한강공원	NA	NA	이쁘/PA	~ ~ ~ ~ ~ ~	1	0
71	여의도한강공원	상푸 광고 촬영... 여의도한강공원	여의도한강 공원	촬영/NC	NA	NA	0	0
72	여의도한강공원	마치 우리 사이처럼 하늘도 맑아 한강 여의도 공원	한강 여의도 공원	맑/PA	NA	NA	1	0
73	여의도한강공원	오랜만에 한강에 왔다 한강 여의도 한강공원 날씨 좋다 한강 여의도 공원	한강 여의도 공원	좋/PA	NA	NA	1	0
74	여의도한강공원	가을라이딩 자전거타기 이동수단의 방편이지만... 그래도 가을햇살 좋은 한강라이딩은 좋다 안녕하신가영 님의 노래는 이런분위기에도 좋아요 한강 여의도 공원	한강 여의도 공원	좋/PA	NA	NA	1	0
75	여의도한강공원	한강치킨♥ 내가 넘 먹고싶었던 한강치킨? ㅎㅎㅎ 한강 치킨 좋아 한국 여행 한강 여의도 공원	한강 여의도 공원	좋/PA	NA	NA	1	0
76	여의도한강공원	한강 한강공원 여의도한강공원 한강캠핑 캠핑 가을이 오는 길목에서 첫 캠핑 바람도 서늘하고 별도 적당하고 참 좋다 ^^♥ 여의도한강시민공원	여의도한강 시민공원	좋/PA	좋/PA	^^♥	2	0
77	여의도한강공원	이쁘다? 여의나루 한강공원	여의나루 한강공원	이쁘/PA	NA	NA	1	0
78	영등포공원	오지 않을 듯 하던 가을이 섬큼 지나가다... 영등포공원	NA	NA	지나/PV	...	1	0
79	오목공원	날씨 좋네!! 오목공원	오목공원	좋/PA	NA	NA	1	0
80	올림픽공원	아침웨이트후 날씨가 너무너무 좋아서 !!! 하남에서~올림픽공원에서~ 마침으로 ~~~건기~~♥♥ 펜잘아 아길레스건염 이야^^ 죽지않을거야 ㅋㅋ 유산소관왕 올림픽공원	올림픽공원, 올림픽공원	건/PV, 좋/PA	건/PV	♥ ~ ~ ♥ ♥ ~ ~	2	0
81	올림픽공원	오빠들 짱 ~ ~ ~ ~ !!! ? 서울 올림픽공원 체조경기장	올림픽공원	짱/NC	짱/NC	~ ~ ~ ~ ~	2	0
82	올림픽공원	공복유산소 운동 올림픽공원 걷기 계단 아작은 무리다...ㅠ ㅠ 일단 걷자... 올림픽공원	올림픽공원, 올림픽공원	건/PV, 운동/NC	NA	NA	1	0

번호	POI	불용어 제거된 트위터 텍스트	공간 감성어 사전		이모티콘 고려		PT	NT
			주어	감성어	공간 감성어 및 추가 감성어	이모 티콘		
83	올림픽공원	오랜만이에요기다렸어요...?? 슈퍼주니어...콘서트 예성 규현 려욱 오랜만이에요 달달한노래 달콤하다 역시 노래쟁 ??? 올림픽공원 올림픽홀	올림픽공원	짱/NC	NA	NA	1	0
84	올림픽공원	토요일 올림픽공원 나들이 비스트 공연 해서 사람 짱 많았다 담당 올만에 셀피 올림픽공원	올림픽공원	나들이/NC	NA	NA	1	0
85	올림픽공원	한국에서 봐서 더 좋았어 올림픽공원 올림픽체조경기장	올림픽공원	쵸/PA	NA	NA	1	0
86	올림픽공원	드디어 본 뮤즈.. 뮤즈 듣고왔던 노래 다 듣고.. 오랜만에 맘홀리며 뛰고.. 즐거웠어요? 올림픽공원 체조경기장	올림픽공원, 올림픽공원	뛰/PV, 즐겁/PA	NA	NA	1	0
87	올림픽공원	즐거운 일요일? 올림픽공원 ()	올림픽공원	즐겁/PA	NA	NA	1	0
88	올림픽공원	첫단독콘서트축하해 루나 선영이 짱짱? 완전멋졌어 올림픽공원 올림픽홀	올림픽공원	짱/NC	NA	NA	1	0
89	올림픽공원	: 지금껏 가장맘에 드는 스텝중. 고생하셨어요 윤화짱. 올림픽공원 올림픽홀	올림픽공원	짱/NC	NA	NA	1	0
90	올림픽공원	다른 구역이지만 사랑스러운 언니랑 비에이피 추억을 또 같이 만들어서 기분이 엄청 좋아요 ㅋㅋㅋ 서울올림픽공원	올림픽공원	쵸/PA	쵸/PA	ㅋㅋㅋ	1	0
91	올림픽공원	..즐겁게??? 서울 올림픽공원 체조경기장	올림픽공원	즐겁/PA	NA	NA	1	0
92	올림픽공원	오늘 콘서트도 대박!!! 내년엔 만나요!!! 아쉬움 신화 신화콘서트 구역 올림픽공원 체조경기장 신화콘서트	NA	NA	대박/NC	!!!!	1	0
93	올림픽공원	날 잡고 싸댕기러 나온 날 ㅋㅋ 어슬렁어슬렁 올팍 산책중, 그 중 요즘 핫플레이스가 된 올림픽공원 들꽃마루 (올림픽 공원 경 중 하나) ~ 일진 서울시 계정으로...	NA	NA	날/NC	ㅋㅋ	1	0
94	올림픽공원	올림픽공원 수영장 다녀오다가 하늘 이 맑아 찍어본 하늘사진 날씨가 따뜻해서 그런지 수영하러 나온 사람들이 많다~ 일만에 해서 그런지 힘드네 올림픽공원 (...	NA	NA	맑/PA	~	1	0
95	올림픽공원	미세먼지 라도 괜찮아 풀린 날씨에 고베 풀린 듯 집나오기 올림픽공원 낙엽길 ??만들어 봤네 ㅋㅋ 가을 정취 물씬 느끼라고 폭신폭신킨 동네 강아지들 좋아할듯...	NA	NA	쵸/PA	...	1	0
96	올림픽공원	이제 년 남았다... 두렵다... 사진은 자변달 내 새끼랑 올림픽공원?? 셀카...	NA	NA	남/PV	...	1	0
97	올림픽공원	아이유 콘서트ㅎ 역시 콘서트는 좋아~ 다른 콘서트들도 좋지만 역시 아이유지ㅋㅋㅋㅋ 아이유 올림픽홀 콘서트 아이유콘서트 올림픽공원 올림픽홀	NA	NA	쵸/PA	~	1	0

번호	POI	불용어 제거된 트위터 텍스트	공간 감성어 사전		이모티콘 고려		PT	NT
			주어	감성어	공간 감성어 및 추가 감성어	이모 티콘		
98	올림픽공원	막내 어린이집 숙제~ 올림픽공원가서 인증사진 찍기!!ㅎㅎ 그런데 막내가 없네~~^^ 막내는 엄마와 뒤에 있고~~ 감만에 올림픽공원 나들이 했네~~!	NA	NA	하/PA	~~	1	0
99	올림픽공원	오늘밤은무슨밤 지누선밤 초호화게스트 벡투터 무대바로앞자리 연예인구경 하겠네~~ 감사 켄터파크 ??? 올림픽공원 올림픽홀	NA	NA	하/PV	~~	1	0
100	올림픽공원	콘서트 보러~~ (올림픽공원 올림픽홀 () 서울특별시. 대한민국)	NA	NA	보/PV	~~	1	0
101	올림픽공원	업그레이드..... 너무 앞인거 아네요??? 괜찮아요 앙?ㅋㅋㅋㅋㅋㅋㅋ 올림픽공원 올림픽홀	올림픽공원	괜찮/PA	NA	NA	0.9812 32	0
102	올림픽공원	마룬파이프 대기중ㅎ 앞에 더티롭스 대박! 올림픽공원 체조경기장	올림픽공원	대박/NC	NA	NA	0.98	0
103	올림픽공원	. 가을 느끼기.. 올림픽공원 소마미술관	올림픽공원	느끼/PV	NA	NA	0	0
104	용마폭포공원	인공잔상이 있는지는 몰랐구만...남양주체육센터에서만 봤는데...그런데 역시나 높다 높아 (용마폭포공원 ())	용마폭포공원	높/PA	모르/PV	...	0	-1
105	월드컵공원	멋진가을날 아빠와아들 이준인 무등이 좋아요~~ 주말 나들이 육아 공원에서 골리기 ㅋㅋㅋ 날씨좋은 월드컵공원	나들이/NC	좋/PA	좋/PA	~~	2	0
106	월드컵공원	. 퇴근후 연남동 에서 저녁 차마시며 소소한애기 상암 에서 영화보려다 그냥 산책 . 미안 .. 상암월드컵공원	월드컵공원	산책/NC	NA	NA	1	0
107	월드컵공원	기분좋게 추운밤날씨 (월드컵공원 () 서울특별시.)	월드컵공원	좋/PA	NA	NA	1	0
108	월드컵공원	아빠바라기-아빠 오는걸 바라보고 있다가 월드컵공원	월드컵공원	바라보/PV	NA	NA	0	0
109	탐골공원	매일 느끼지만... 호선은 탐골공원찍으러 가는 노인들만 없어도 훨씬 편해질것같다.(임산부를 못본채하며 끝까지 앉아있는 파란치들을 본 이후	탐골공원	느끼/PV	NA	NA	0	0
110	파리공원예쁜 나무.. 파리공원	파리공원	예쁘/PA	NA	NA	1	0
111	학동공원	셀스타그램 학동 심심 졸립다 학동공원	학동공원	심심/NC	NA	NA	0	-1
112	학동공원	해맑은 우리 학동공원	학동공원	맑/PA	NA	NA	1	0
113	허브천문공원	허브천문공원 아담하니 좋네요~! 허브천문공원	허브천문공원, 허브천문공원	아담/NC, 좋/PA	좋/PA	~	2.9655 17	0
114	효창공원	아침공기 상쾌 +-+ 효창공원	효창공원	상쾌/NC	상쾌/NC	+_-+	1.9928 57	0
115	효창공원	퇴근(?)하고 취미 활동 하러 ~ (효창공원앞 번 출구)	효창공원	활동/NC	하/PV	~	1	0

<부록 D> 감성분석 결과에 대한 설문조사 응답 결과

서울 테마공원에 대한 트위터 텍스트 감성분석 서베이입니다.

안녕하세요. 본 설문에 응해주셔서 감사합니다.)

본 설문은 서울에 위치한 테마공원에 대해 트위터 텍스트에 나타난 감성을 묻는 설문입니다.

109개의 질문이 있으며, 10분 정도 소요됩니다.

설문을 마친 후 마지막 응답에 핸드폰 번호를 남겨주시면 추첨을 통해 기프티콘을 보내드립니다~!

각 질문의 트위터 텍스트를 읽고 장소에 대한 트위터 작성자의 감성을 선택해주세요.

예1) 낙산공원은 역시 야경이 예뻐네 : '긍정' 선택

예2) 롯데월드 사람 많아서 돌아다니기 너무 힘들다 ππππ : '부정' 선택

1. #셀스타그램 #학동 #심심 #줄립다 @ 학동공원

☐ 긍정

☐ 중립

☐ 부정

2. 얼마나 고마운지 모르겠어! 아무리 사진이 별로 좋지 않아도 가장 좋은 추억을 남겼다~ 히힃 #카우 #제인 #하나 #추억이라부르게 @ 여의도한강공원

☐ 긍정

☐ 중립

☐ 부정

3. 예고예고 땀 줄줄 힘들다. 여기서 여의도 가로지르면 집인데 그래도 안양천까지 가서 도림천타고 들어갈거다. 내일은 지각 확정이겠군 여의도한강공원

☐ 긍정

☐ 중립

☐ 부정

그림 D-1. 감성분석 설문지 일부

표 D-1. 감성분석 결과에 대한 설문조사 응답 결과

문항 번호	트위터 텍스트	긍정 응답 비율	중립 응답 비율	부정 응답 비율	설문 결과 (Act ual)	감성분석 결과 (Predict ed)
1	1. #셀스타그램 #학동 #심심 #졸립다 @ 학동공원	7%	54%	38%	중립	부정
2	2. 얼마나 고마운지 모르겠어! 아무리 사진이 별로 좋지 않아도 가장 좋은 추억을 남겼다~ 허형 #카우 #제인 #하나 #추억이라부르게 @ 여의도한강공원	80%	5%	15%	긍정	부정
3	3. 에고에고 땀 줄줄 힘들다. 여기서 여의도 가로지르면 집인데 그래도 안양천까지 가서 도립천타고 들어갈거다. 내일은 지각 확정이었군 여의도한강공원	12%	41%	47%	부정	부정
4	4. 환상에 나라 롯데월드 힘들다 #롯데월드 #lotteworld #잠실 #놀이기구 #환상에나라 #꿈과희망 #어지러움	11%	37%	52%	부정	부정
5	5. 퇴근후 날아와서 잠실 이성당 #이성당 #잠실 #9호선 #힘들어 @ 롯데월드	15%	44%	41%	중립	부정
6	6. #캘리포니아피자키친 #BBQ피자 #올ㅋ존맛ㅋ #빵이마이?! #매장이겁나한가해서그런가(...)서버분이엄청리필여부물어 봐주셨다(...) @ 캘리포니아 피자키친 롯데월드몰점	52%	25%	23%	긍정	부정
7	7. 왜 나는 이런 날씨에만 녹화를 나오는가... 태안에서 눈보라 맞은 이후 정확히 1년후 상암동 겨울왕국을 만나다. #majorsnowday #fml #himalaya @ 노을공원	7%	31%	62%	부정	부정
8	8. 흐린 날씨... #돈키우스 #트레킹 #노을공원 #Landscape @ 노을공원	11%	49%	40%	중립	부정
9	9. 인공암장이 있는지는 몰랐구만...남양주체육센터에서만 봤는데...그런데 역시나 높다 높아 @ 용마록포공원	14%	46%	40%	중립	부정
10	10. 둘이서만 와보는 #롯데월드 이게 얼마만인가 ㅋ 아틀란티스 점검인게 아쉽군... @ 환상과모험의나라 롯데월드	47%	32%	21%	긍정	부정
11	11. 롯데월드 한동안 못가 느~~~~무 힘들어 ㅋㅋㅋ (마지막쪽에 롯데월드 댄서들과 배경음악이 딱맞아ㅋㅋㅋ) #잠실 #롯데월드 #꿈과희망의나라 #lotteworld #페스티벌...	25%	46%	30%	중립	부정
12	12. 사슴아 침은 묻히지 말고 먹으면 안되겠니... #어린이대공원 #동물원 #먹이주기 #사슴 @ 어린이대공원	49%	35%	16%	긍정	부정
13	13. 월드시네마에서 #서부전선 감상.. 중간중간은 한참 웃었는데... 내용은 다 예상되지만 생각없이 웃을수 있었네.. 항상 그렇지만 끝은 아쉬웠네... @ 잠실 롯데월드시네마	28%	48%	23%	중립	부정
14	14. 허브천문공원 아담하니 좋네요~! @ 허브천문공원	91%	4%	5%	긍정	긍정
15	15. 실물 처음 봄 ㅋㅋㅋㅋㅋㅋㅋㅋ 신기하게 생겼다 ㅋㅋㅋ @ 롯데월드몰	68%	30%	2%	긍정	긍정
16	16. 간단에 날씨 좋아서 아들 데리고 나들이. (@ 어린이대공원 (Children's Grand Park) in 서울특별시.)	80%	16%	4%	긍정	긍정
17	17. #멋진가을날 #아빠와아들 #lovelyjoons 이준인 무등이 좋아요~~~ #주말 #나들이 #육아 #공원에서 #글리기 ㅋㅋㅋ @ 날씨좋은 월드컵공원	86%	12%	1%	긍정	긍정
18	18. 아이 좋아~ @ 노을공원 캠핑장	77%	22%	1%	긍정	긍정
19	19. 아주 예쁘다~~~ #꽃전구 #flower #flowerbulbs @ 동대문역사문화공원 DDP	84%	15%	1%	긍정	긍정
20	20. 좋다~ (@ 여의도한강공원 (Yeouido Hangang Park) - @seoulhangang in 서울특별시, 서울시)	86%	10%	4%	긍정	긍정

문항 번호	트위터 텍스트	긍정 응답 비율	중립 응답 비율	부정 응답 비율	설문 결과 (Act ual)	감성분석 결과 (Predict ed)
21	21. 아침웨이트후 날씨가 너무너무 좋아서 !!! 하남에서~올림픽공원에서~~마침으로 ~~건기~~♥♥괜찮아 #아킬레스건염 이야^^ 죽지않을거야 ㅋㅋ#유산소끝판왕 @ 올림픽공원	77%	16%	7%	긍정	긍정
22	22. 2015 god Concert 오빠들 찡~~~!!! @ 서울 올림픽공원 제조경기장	72%	26%	2%	긍정	긍정
23	23. 잠실의 메카 롯데월드2... 멋지기는 하네... 안개가 아주짱 @ 롯데월드 (lotteworld)	68%	25%	7%	긍정	긍정
24	24. 가족 나들이~ 어린이대공원에서 내식구 총출동~!!! (@ 어린이대공원 동물원 (Children's Grand Park Zoo) in 서울특별시.)	79%	20%	1%	긍정	긍정
25	25. 아침공기 상쾌 ++ ????? @ 효창공원	73%	25%	2%	긍정	긍정
26	26. 산책 (@ 도산공원 (Dosan Park) in 서울특별시, South Korea)	52%	46%	2%	긍정	긍정
27	27. #멜팅샵#meltingshop#현대카드#고메위크#2015 #lovebasil..#바질 사랑하여 ㅋㅋㅋ트트트 #소프트셀크랩#카프레제파스타 #장백기 왔다 ㅋㅋㅋ트 @ 도산공원...	44%	52%	4%	중립	긍정
28	28. 오늘은 회사언리랑 도산공원까지?? 단풍구경~!!?ㅎㅎ #운동 #걷 #산책 #압구정 #도산공원 #점심시간 #쉬는시간 #행복 #단풍 #가을 #韓? #休み @ 도산공원	60%	36%	3%	긍정	긍정
29	29. 도산공원 산책하다 #도산공원 #서울 #강남 #seoul #gangnam #dosanpark #ソウル #江南 #島山公園 #首爾 #首爾 @ 도산공원	57%	41%	2%	긍정	긍정
30	30. 헤헤헤 여긴 일부러 이렇게 레이어로 쌓아주는걸까 암튼 넘넝 맛이쑈 우린 아리아개는 언제 가보는걸까? ㅋㅋㅋ #런치 #오마카세 #스시 #스시선수 #도산공원 #도산공원맛집ㅋㅋㅋ	69%	26%	5%	긍정	긍정
31	31. . #퇴근후 #연남동 에서 #저녁 차마시며 #소소한얘기 #상암 에서 영화보려다 그냥 #산책 ? . . #미안 .. @ 상암월드컵공원	49%	38%	12%	긍정	긍정
32	32. 기분 좋게 추운밤날씨 (@ 월드컵공원 (Seoul World Cup Park) in 서울특별시.)	69%	25%	6%	긍정	긍정
33	33. ..pretty trees..예쁜 나무.. @ 파리공원	72%	25%	4%	긍정	긍정
34	34. 급 번개 나들이 :) (@ 노을공원 in 서울특별시.)	65%	31%	4%	긍정	긍정
35	35. 동대문 롯데퍼트인 컬쳐스타에 레이저수잔의 공간이 생겼어요!!) 완벽하진않지만 이번주내로 야금야금 채워나갈게요. 동대문역사문화공원역 11-12번출구 사이에 바로 들어올수있는...	46%	51%	4%	중립	긍정
36	36. 해맑은 우리 #prame @ 학동공원	59%	37%	4%	긍정	긍정
37	37. 날씨 좋네!! #sbs #mokdong @ 오목공원	67%	32%	1%	긍정	긍정
38	38. 도서관도 있고 좋네 (@ 낙성대공원 (Nakseongdae Park) in 서울특별시.)	77%	23%	0%	긍정	긍정
39	39. 짐사람이랑 밖에서 저녁 먹고 산책~ㅋ 피톤치드 누려~!!ㅋ (@ 북서울꿈의숲 (Dream Forest) - @adidasrunningkr in Seoul)	68%	32%	0%	긍정	긍정
40	40. 북서울꿈의숲 전망대 집에 오고서야 불꽃축제가 있었다는 걸 알게됨 알았으면 저 위에서 좀 더 기다려볼 걸 그래도 안보이려나 ㅋ @ 북서울 꿈의 숲	53%	40%	7%	긍정	긍정

문항 번호	트위터 텍스트	긍정 응답 비율	중립 응답 비율	부정 응답 비율	설문 결과 (Act ual)	감성분석 결과 (Predict ed)
41	41. 오늘 서울 하늘 맑음 #여직원회식#여의도#한강#나들이 @ 여의도한강공원	72%	26%	2%	긍정	긍정
42	42. 빼에로 아줌마. 몇년 전부터 주로 서울광장에서 보이시는 아주머니 간혹 서울숲 여의도한강공원 남산 등지에서도 보이던... (복장 + 얼굴칠 은 항상 동일 빨간코는...	33%	54%	12%	중립	긍정
43	43. 서강대교 아래 트랙 한바퀴 윙리로 돌기 성공~ ~!#라이딩 #자전거 #파이어아이 #fireeyebike @ 여의도한강공원 서강대교 아래	68%	31%	1%	긍정	긍정
44	44. #공복유산소 #운동 #올림픽공원 #걷기 #계단 아직은 무리다...ㅠㅠ 일단 걷자... @ 올림픽공원	37%	49%	14%	중립	긍정
45	45. RT @SsoSso_1004: 오랜만에요기다렸어요K.R.Y?? #슈퍼주니어#K.R.Y#콘서트#예성#규현#려욱#오랜만에요#달달한노래#달콤하다#역시#노래짱 ??? @ 올림픽공원 올림픽홀	59%	33%	7%	긍정	긍정
46	46. #토요일#올림픽공원#나들이#비스트#공연#해서#사람#짱#많았다#덤당#올만에#셀피 @ 올림픽공원	49%	28%	22%	긍정	긍정
47	47. 한국에서 봐서 더 좋았어??#b2st #beautifulshow2015 @ 올림픽공원 올림픽체조경기장	65%	25%	10%	긍정	긍정
48	48. 드디어 본 뮤즈.. #뮤즈 듣고왔던 노래 다 듣고.. 오랜만에 맘홀리며 뒹고.. 즐거웠어요? @ 올림픽공원 체조경기장	72%	26%	2%	긍정	긍정
49	49. 즐거운 일요일? @ 올림픽공원 (Olympic park)	67%	31%	2%	긍정	긍정
50	50. #fx #첫단독콘서트축하해 #루나 #선영이 #짱짱? #완전멋졌어 @ 올림픽공원 올림픽홀	65%	30%	5%	긍정	긍정
51	51. RT @Famous_Jae: 지금껏 가장만에 드는 스탭증. 고생하셨습니다 윤화짱. #fx #Dimension4 #DockingStation @ 올림픽공원 올림픽홀	54%	36%	10%	긍정	긍정
52	52. 다른 구역이지만 사랑스러운 @renlvtop 언니랑 비에이피 추억을 또 같이 만들어서 기분이 엄청 좋아요 ㅋㅋㅋ @ 서울올림픽공원	79%	15%	6%	긍정	긍정
53	53. Last..즐겁게??? @ 서울 올림픽공원 체조경기장	67%	27%	6%	긍정	긍정
54	54. 운이 좋았어!! @ 제2롯데월드몰	69%	28%	2%	긍정	긍정
55	55. 요다 ㅋ 롯데월드몰 롯데시네마에서 처음으로 영화봄 ㅋ 다른건 다 좋은데 사운드가 병맛 (심지어 사운드 특화관인 슈퍼사운드? 관이었는데...) 입체감 전혀 없이 앞에서만 나와서...	46%	11%	43%	긍정	긍정
56	56. 지금 집에서 아침먹으며 보는 하늘이 꽤 아름답네 #seoul #lotteworld #lotteworldtower #sky #cloud #jamsil @ 롯데월드 (lotteworld)	74%	25%	1%	긍정	긍정
57	57. #에스타그램 #셀스타그램 #팔스타그램 #육스타그램 #소통 #육스타그램 #육아#베이비그램 #베베 #베이비인스타 #6개월 #롯데월드몰#맛팔 #좋아요 #선팔 #베이비 #줍스타그램 #육아...	38%	53%	9%	중립	긍정
58	58. : #20151002 #금요일 #서울 #잠실 #롯데월드 #모험 파 신비가 가득한 나라 우리가 꿈꾸던 그곳 여기는~ 롯데월드~♡ #바이킹 이랑 #자이로드롭 은...	68%	25%	7%	긍정	긍정
59	59. #marimekko #socks #unikko #마리메꼬 #양말 #우니꼬 우니꼬 양말은 언제나 기분이 좋아??????? @ 롯데월드몰	63%	31%	6%	긍정	긍정

문항 번호	트위터 텍스트	긍정 응답 비율	중립 응답 비율	부정 응답 비율	설문 결과 (Actual)	감성분석 결과 (Predict ed)
60	60. 롯데월드 세계 최대의 실내 놀이공원 #롯데 #롯데월드 #놀이공원 #서울 #한국 #대한민국 #아이스링크 #웃겨 #좋아 #사진 #사진스타그램 #테일리 #맛팔 Lotte World...	64%	28%	7%	긍정	긍정
61	61. #힐링나들이 #서울 #seoul #lotteworld #롯데월드 #꿈과희망보다는 #비명과광란 #사진은많은데 #폐북에올려버림 #다들고마워용 #? @ 롯데월드 (lotteworld)	62%	16%	22%	긍정	긍정
62	62. 자본주의 최고야 (@ 롯데월드몰 (LOTTE WORLD MALL) in Seoul)	62%	26%	12%	긍정	긍정
63	63. 151213. Sun. #롯데월드 가든스테이지 깨끗하게! 말개! 자신있게! ? 미소천사 ? 남시오 ~ exsy89 @jamesjoohyunlee @moonchul...	68%	27%	5%	긍정	긍정
64	64. 내일 시험이 있는데..... #롯데월드 #크리스마스 #분위기 #너무 #좋다 ?? @ 롯데월드 (lotteworld)	67%	28%	5%	긍정	긍정
65	65. Coffee + plants ?? + ? . #cafe #coffee #coolplace #seoul #gangnam #도산공원 #카페 . 한적할때 더 좋은...	67%	28%	5%	긍정	긍정
66	66. 2015. 12. 10. #도산공원 앞 #커피지인 #도피오 #에스프레소터블 오랫동안 진한 에스프레소 맛나네... @ Coffee Jiin	62%	36%	2%	긍정	긍정
67	67. 퇴근(?)하고 취미 활동 하러 ~ (@ 효창공원앞 4번 출구)	44%	56%	0%	중립	긍정
68	68. . 내일은 우리 뽕&찹쌀군 데리고 산책해야지~ . #산책 #공원 #매미소리좋다 #강아지 #개 #남산공원 #park #walk #stroll #dog @ 남산공원 Namsan Park	59%	35%	6%	긍정	긍정
69	69. #남산공원 처음으로 갔다. 길 몰라서 재미있었다 ㅋㅋ...	64%	33%	2%	긍정	긍정
70	70. 좀 이쁘긴 하더라~ #LED #장미 #꽃밭 #동대문역사문화공원 #DDP @ 동대문 디자인 플라자 Led 장미 꽃밭	81%	17%	1%	긍정	긍정
71	71. 시작 ... (@ 낙성대공원 (Nakseongdae Park) in 서울특별시.)	40%	53%	7%	중립	긍정
72	72. 오지 않을 듯 하던 가을이 섬کم 지나가다... #korea #seoul #seoulis #Autumn #iphone6 @ 영등포공원	42%	54%	4%	중립	긍정
73	73. 150723 예비둘째형부(?)랑 둘째언니랑 엄마랑 함께... 두번째 #뮤지컬문준경 보러 ㅎㅎ (@ 북서울꿈의숲 아트센터 (Dream Forest Art Center) in 서울특별시. South Korea)	67%	32%	1%	긍정	긍정
74	74. 이쁘다~ ~ ~ ~ ~ @ 바람살랑살랑부는 여의도한강공원	78%	20%	2%	긍정	긍정
75	75. 오늘 콘서트도 대박!!!! 내년엔 만나요!!!! #아쉬움 #신화 #신화콘서트 #F6구역 @ 올림픽공원 제2경기장 신화콘서트	58%	30%	12%	긍정	긍정
76	76. 날 잡고 싸매기러 나온 날 ㅋㅋ 어슬렁어슬렁 울퉁 산책중. 그 중 요즘 핫플레이스가 된 올림픽공원 들꽃마루 (올림픽 공원 9경 중 하나) 2~3일전 서울시 SNS계정으로...	74%	23%	2%	긍정	긍정
77	77. #올림픽공원 #수영장 다녀오다가 #하늘 이 맑아 찍어본 #하늘사진 날씨가 따뜻해서 그런지 수영하러 나온 사람들이 많다~ 3일만에 해서 그런지 힘드네 @ 올림픽공원 (Olympic...	79%	19%	2%	긍정	긍정
78	78. #미세먼지 라도 괜찮아 풀린 날씨에 고삐 풀린 듯 걸나오기 #올림픽공원 #낙엽길 ??만들어 냈네 ㅋㅋ #가을 정취 불문 느끼라고 폭신폭신 동네 강아지를 좋아 할듯...	67%	27%	6%	긍정	긍정

문항 번호	트위터 텍스트	긍정 응답 비율	중립 응답 비율	부정 응답 비율	설문 결과 (Actual)	감성분석 결과 (Predicted)
79	79. 이제 1년 남았다... 두렵다... 사진은 저번달 내 새끼랑 올림픽공원?? #selfie #dog #puppy #poodle #blackpoodle #toypoodle #셀카...	27%	67%	6%	중립	긍정
80	80. 아이유 콘서트ㅎ 역시 콘서트는 좋아~ 다른 콘서트들도 좋지만 역시 아이유지ㅋㅋㅋㅋ #아이유 #올림픽홀 #콘서트 #아이유콘서트 @ 올림픽공원 올림픽홀	67%	32%	1%	긍정	긍정
81	81. 막내 어린이집 숙제~ 올림픽공원가서 인증사진 찍기!!ㅎㅎ 그런데 막내가 없네~^^ 막내는 엄마와 뒤에 있고~~ 감만에 올림픽공원 나들이 했네~~! @ Olympic Park Seoul	62%	37%	1%	긍정	긍정
82	82. #오늘밤은무슨밤 #지누션밤 #초호화게스트 #백투더90s #무대바로앞자리 #연예인구경 하겠네~~#감사 #전터파크 ??? @ 올림픽공원 올림픽홀	67%	28%	5%	긍정	긍정
83	83. MC THE MAX 콘서트 보러~~ (@ 올림픽공원 올림픽홀 (Olympic Park Olympic Hall) in 서울특별시. 대한민국)	44%	53%	2%	중립	긍정
84	84. 지난 8월 마지막에 먹은 #롯데월드 입장전에 먹은 아이스크림^^ 맛났다♥? #먹스타그램 #맛스타그램 #food #instafood #lotteworld #icecream #ice...	62%	36%	2%	긍정	긍정
85	85. 난 여기 야채빵 단팥빵보다 크림 든 빵들이 맛있더라~ #이성당 @ 롯데월드 (lotteworld)	62%	37%	1%	긍정	긍정
86	86. #제미난놀이 #부메랑 #옴짤 #운동화 #롯데월드점 참 재미난게 많네 ㅎㅎ @ Lotte World	68%	30%	2%	긍정	긍정
87	87. EXO SingForYou 쇼케이스 ?? 생각보다 완전 가깝다!!ㅠㅠ 얼른 보고싶어 사랑해 ?? #EXO #SingForYou #쇼케이스 #롯데월드 @ 롯데월드 (lotteworld)	62%	32%	6%	긍정	긍정
88	88. 151213. Sun. #롯데월드 가든스테이지 그댄 먼 곳만 보네요~ 내가 바로 여기 있는데... (러브홀릭 '인형의 꿈' 중에서) @jamesjoohyunlee...	38%	56%	6%	중립	긍정
89	89. 지난 방학때 서울 롯데월드에 가고 친구들과 함께 즐겁게 놀았다~ (Went to Lotte World in Seoul at last holiday with my...	74%	22%	4%	긍정	긍정
90	90. : #20151002 #금요일 #서울 #잠실 #롯데월드 완전 멋있었던 야간 #퍼레이드... #놀이동산 #로터 #로리 #럽스타그램 #신밋드의모험 #후렌치레볼루션...	77%	17%	6%	긍정	긍정
91	91. 밝은녀석 @ 롯데월드 (lotteworld)	54%	42%	4%	긍정	긍정
92	92. 으악야아 넓어 (@ 롯데월드몰 (LOTTE WORLD MALL) in Seoul)	52%	41%	7%	긍정	긍정
93	93. 업그레이드..... 너무 앞인거 아네요??? #괜찮아요 잉?ㅋㅋㅋㅋㅋㅋㅋㅋ @ 올림픽공원 올림픽홀	53%	37%	10%	긍정	긍정
94	94. 드디어! 근데 생각보다 괜찮음ㅋ (@ 롯데월드몰 (LOTTE WORLD MALL) in Seoul)	70%	27%	2%	긍정	긍정
95	95. 바지락 매생이 비빔면 대박! #멜팅샵 @ 도산공원 Melting Shop	63%	35%	2%	긍정	긍정
96	96. 마루파이크 대기중ㅎ 앞에 더티롭스 대박! @ 올림픽공원 체조경기장	60%	30%	10%	긍정	긍정
97	97. 고 투더 헬!!!! 영혼의 안식처..... 모즈 신사동!! #모즈#신사동모즈#일식#샷포#첫번째#?사이모즈#신사동#도 산공원#해본#천국의 창고 ㅎ ㅎ @ ?사이 모즈	59%	35%	6%	긍정	긍정
98	98. 비오기전 #단풍 구경 #몽마르뜨공원 #상쾌한공기 #토끼천국 @ 서레마을 몽마르뜨 공원	74%	23%	2%	긍정	긍정

문항 번호	트위터 텍스트	긍정 응답 비율	중립 응답 비율	부정 응답 비율	설문 결과 (Act ual)	감성분석 결과 (Predict ed)
99	99. #생일쿠폰 으로 내가 참 좋아하는 #티라미수 #로열자몽티 홀짝. 역시 자몽티는 꽤 진하게 잘 만들어졌어. 조용하다 #일요일저녁 #도산공원 #아티제 @아티제 도산공원점	65%	27%	7%	긍정	중립
100	100. 9월 3일 종료. 오늘 하루 종일 두통때문에 엄청 힘들었는데 마지막은 큰 선물을 받는 느낌. #롯데월드몰 #길리안 #초콜릿 #카페 #초콜릿빙수 이런 #빙수 처음이야...! 달아...	49%	43%	7%	긍정	중립
101	101. #제2롯데월드 높다.? #일상 #DAILY #빌딩 #팔로우 #선팔 #맛팔 #맛팔해요 #선팔해요 #인친 #잠실 #송파구 #서울 #석촌호수 #좋아요 @ 제2롯데월드몰	37%	51%	12%	중립	중립
102	102. 오늘의 점심은 높은데서. #도림 #Chinese #lunch #lottehotel @ 롯데월드호텔-도림	36%	58%	6%	중립	중립
103	103. 이곳에서도 보일 정도로 #제2롯데월드 가 높아졌다. 어릴 때에는 뉴욕 엠파이어와 63빌딩이 최고인 줄만 알았는데, 급성장한 중국에는 저런 게 수백 동이나 된다고 하고... @...	37%	49%	14%	중립	중립
104	104. 열심히 걷고 있었음 (@ 롯데월드 아쿠아리움 (Lotte World Aquarium) in 서울)	22%	68%	10%	중립	중립
105	105. 이런 WTF!!! 팀 쿡이 잡스?의 유산 #애플 을 "삼성 따라쟁이"로 만들었다 시바... S펜 하고 #애플펜 하고 다르게 뭐임? (@ 서서울호수공원 in 서울특별시, Seoul 서울특별시.)	6%	42%	52%	부정	중립
106	106. 삼푸 광고 촬영... @ 여의도한강공원	19%	74%	7%	중립	중립
107	107. #가을 느끼기.. #올림픽공원 #소마미술관 @ Olympic Park Seoul	38%	58%	4%	중립	중립
108	108. 아빠바라기-아빠 오는걸 바라보고 있다가 @ 월드컵공원	27%	70%	2%	중립	중립
109	109. 매일 느끼지만...1호선은 탑골공원쪽으러 가는 노인들만 없어도 훨씬 편해질것같다.(임산부를 못본채하며 끝까지 앉아있는 파렴치들을 본 이후	11%	22%	67%	부정	중립
110	110. 마치 우리 사이처럼 하늘도 맑아 @ 한강 여의도 공원	59%	40%	1%	긍정	긍정
111	111. 오랜만에 한강에 왔다 #한강 #여의도 #한강공원 #날씨 #좋다 @ 한강 여의도 공원	54%	43%	2%	긍정	긍정
112	112. #가을라이딩 #자전거타기 #이동수단 의 방편이지만... 그래도 #가을햇살 좋은 #한강라이딩 은 #좋다 #안녕하십니까영 님의 노래는 이런분위기에든 #좋아요? @ 한강 여의도 공원	68%	30%	2%	긍정	긍정
113	113. 한강치킨♡ 내가 넘 먹고싶었던 한강치킨?ㅎㅎㅎㅎ #한강 #치킨 #좋아 #한국 #여행 @ 한강 여의도 공원	72%	27%	1%	긍정	긍정
114	114. #한강 #한강공원 #여의도한강공원 #한강캠핑 가을이 오는 길목에서 첫 캠핑 바람도 서늘하고 별도 적당하고 참 좋다 ^^♥ @ 여의도한강시민공원	83%	17%	0%	긍정	긍정
115	115. 이쁘다? @ 여의나루 한강공원	65%	35%	0%	긍정	긍정

A Method for Analysis of Preferences of
Places based on Sentimental Analysis
using SNS Data
: Case Study on Theme Parks in Seoul,
South Korea

Supervised by Prof. Yu, Kiyun

August 2017

Chae, Inyoung

Department of Civil and Environmental Engineering
Seoul National University

Abstract

With the proliferation of mobile devices and the activation of various social networking services, SNS users share their daily lives and build social networks through SNS. Users write experiences and feelings about their place of visits on the social networking site, which is shared by others and affects their decision making. A number of postings are uploaded through SNS, enabling analysis to derive meaningful information from them. Therefore, this study suggests a formula for performing sentimental analysis on SNS text

and quantifying the preferences of places. In order to perform sentimental analysis, this study utilizes the existing spatial sentiment lexicon and considers emoticon as important sentimental information reflecting the characteristics of SNS text. Based on the analysis of preferences of places proposed in this study, the preference of 36 theme parks located in Seoul was analyzed. These results can be usefully utilized in place information services in the future. This study is meaningful in that it can analyze the preferences of places using unstructured text data such as blogs, cafes, and news articles that users write everyday, as well as SNS text.

**Keywords : SNS text, Twitter, Emoticon, Spatial Sentiment
Lexicon, Sentimental Analysis,
preferences of places**

Student Number : 2015-22937